The AMERICAN PHYSICS TEACHER

A Journal Devoted to the Instructional and Cultural Aspects of Physical Science

VOLUME 7 | DECEMBER, 1939 | NUMBER 6

The Propagation of Errors

RAYMOND T. BIRGE
Department of Physics, University of California, Berkeley, California

THE question of what constitutes the most reliable value to be assigned as the uncertainty of any given measured quantity is one that has been discussed for many decades and, presumably, will continue to be discussed. It is a question that involves many considerations and by its very nature has no unique answer. The subject of the propagation of errors, on the contrary, is a purely mathematical matter, with very definite and easily ascertained conclusions. Although the general subject matter of the present article is by no means new,[1] many scientists still fail to avail themselves of the enlightening conclusions that may often thus be reached, while others frequently use the theory incorrectly and thus arrive at quite misleading conclusions.

The specific problem is: Given a definitely assigned uncertainty for each of a set of *independently* measured quantities, what is the resulting uncertainty in any specified *function* of these quantities? If each assigned uncertainty is intended to represent a probable error, the calculated uncertainty is the corresponding probable error of the function. If each assigned uncertainty is a root-mean-square error, the result is a root-mean-square error, etc. The only necessary assumptions are that (1) each error is only a small fraction of the corresponding observed quantity, and (2) positive and negative errors are equally likely to occur. These two conditions are normally fulfilled in scientific work.

Let the independently observed quantities be denoted by $z_1z_2z_3\cdots$. Let these be combined in any way to give the function

$$Z=f(z_1z_2z_3\cdots). \tag{1}$$

Suppose that the true, but necessarily unknown errors of $z_1z_2\cdots$ are $e_1e_2\cdots$, respectively. Then the true, but unknown error in Z is E, where

$$Z+E=f(z_1+e_1,\ z_2+e_2,\ \cdots). \tag{2}$$

To get a simple relation between E and the various e's, we apply the Taylor expansion to Eq. (2). Because of our first assumption that each $e_i/z_i \ll 1$, only first-order terms need be retained. Hence

$$(Z+E)-(Z)=E=(\partial Z/\partial z_1)e_1 + (\partial Z/\partial z_2)e_2+\cdots. \tag{3}$$

We now square both sides of Eq. (3). On the right-hand side we will then have terms containing e_i^2 and other terms containing e_ie_j. Since we have assumed that each e_i is equally likely to be positive or negative, some of the cross-product terms will add to the magnitude of E^2, while others will diminish it. In any such single set of measurements the sign, like the magnitude of each cross-product term, is quite definite, although necessarily unknown.

Our real object, however, is to make a reliable calculation of the uncertainty R in Z, on the

[1] See, for instance, M. Merriman, *Method of Least Squares* (ed. 8), pp. 75–79; A. de F. Palmer, *Theory of Measurements*, pp. 95–104.

basis of *assigned* uncertainties $r_1r_2\cdots$ in $z_1z_2\cdots$. This is accomplished as follows. One writes Eq. (3) for each of a very large number of independent sets of measurements, adds the complete set of equations thus obtained, and divides by n, the number of equations thus summed. The left-hand side of this final result is $\Sigma E^2/n$ and this is just the definition of R^2, where R is the root-mean-square error of Z. On the right-hand side the sum of the cross-product terms will be vanishingly small (since they are equally likely to be positive or negative), compared to the sum of the remaining terms. Hence the cross-product terms may be ignored, and there remain only terms that can be summed into groups of the form $(\partial Z/\partial z_i)^2\Sigma e_i^2/n$. But $\Sigma e_i^2/n$ defines r_i^2, where r_i is the root-mean-square error in z_i. Hence, finally,

$$R^2=(\partial Z/\partial z_1)^2r_1^2+(\partial Z/\partial z_2)^2r_2^2+\cdots, \tag{4}$$

and this equation is known as the *law of propagation of errors.*

A root-mean-square error can always properly be assigned to a measured quantity, regardless of the assumed law of *distribution* of errors, but the *significance* of such an error (for example, the probability of exceeding it, in any single measurement) can be calculated only on the basis of some assumed distribution. Thus, if one assumes a normal (Gaussian) distribution of errors, the probability of exceeding the root-mean-square error is 31.73 percent, and the probable error (for which the corresponding probability is just 50 percent) is 0.6745 times the root-mean-square error. None of these details, however, affects the validity of Eq. (4). If probable errors r_i can properly be assigned, then R in Eq. (4) is the resulting probable error in Z. If only root-mean-square errors r_i can be assigned, then R in Eq. (4) represents the root-mean-square error of Z.

It may be noted parenthetically, at this point, that in connection with the design of apparatus for the evaluation of any quantity Z, which is given by a $f(z_1z_2\cdots)$, the most efficient distribution of errors among the z_i's is that which will make all terms of Eq. (4) of equal magnitude. This is the *principle of equal effects.* There is no real advantage in expending time and money on an attempt to improve further the accuracy of a certain z_i when Eq. (4) shows that the final uncertainty of Z is due mainly to some other z_i. The first object should be to improve, if possible, the accuracy of this latter z_i.

Although Eq. (4) is entirely general, and hence may be applied to any function, it is quite customary, in actual scientific work, to use some special form of the equation, applying to a special form of the function Z. As a matter of fact, most of the functions appearing in ordinary scientific work can be treated more rapidly and conveniently by the use of some appropriate special form of Eq. (4). Unfortunately, however, it is quite easy to apply a special form improperly, and hence one of the purposes of the present paper is to list some of these special forms, and to point out the functions to which they do and do not apply.

It is convenient to denote the ordinary, or absolute, probable error in a measured quantity z_i by r_i, and the *proportional* probable error r_i/z_i by t_i. Similarly R/Z will be denoted by T. We may then derive many special forms of Eq. (4), among which the following are especially useful.

(*a*) If
$$Z=z_1+z_2+\cdots z_n, \tag{5}$$
$$R=(r_1^2+r_2^2+\cdots r_n^2)^{\frac{1}{2}}. \tag{6}$$

Any z_i may be positive or negative, and thus Eq. (6) applies to a function consisting of any series of terms, regardless of sign. In particular it applies to the sum or difference of any two quantities, in which case $n=2$.

(*b*) If
$$Z=(1/n)(z_1+z_2+\cdots z_n), \tag{7}$$
$$R=(1/n)(r_1^2+r_2^2+\cdots r_n^2)^{\frac{1}{2}}. \tag{8}$$

But Eq. (7) represents the simple arithmetic average of n observations z_i and such an average is valid only if each z_i should be assigned the same weight. In such a situation $r_1=r_2=r_3\cdots=r$. Hence

$$R=(1/n)(nr^2)^{\frac{1}{2}}=r/n^{\frac{1}{2}}. \tag{8'}$$

Thus the familiar rule for the probable error of the average of n equally reliable observations follows directly from the general law of propagation of error.

(*c*) Let Z be the weighted average of n observations with weights p_i. Thus

$$Z=\Sigma p_iz_i/\Sigma p_i. \tag{9}$$

If the weights have been properly assigned,

$$r_1 = r/(p_1)^{\frac{1}{2}}, \quad r_2 = r/(p_2)^{\frac{1}{2}}, \quad \text{etc.,} \tag{10}$$

where r is the probable error of a hypothetical observation of unit weight. Then Eq. (4) yields

$$R = r/(\Sigma p_i)^{\frac{1}{2}}, \tag{11}$$

which is the familiar equation for the probable error of a weighted average.

(*d*) Let $$Z = z^n. \tag{12}$$

Then, from Eq. (4),

$$R/Z = nr/z \quad \text{or} \quad T = nt. \tag{13}$$

The proportional probable error in z^n is thus n times the proportional probable error in z. This holds for *all* values of n, including proper and improper fractions. Furthermore, n in Eq. (12) may equally well be negative. In this connection it may be noted that n in Eq. (13) has no sign, since all symbols for errors, such as R and T, represent *magnitudes* only. Eq. (4) yields only intrinsically positive quantities such as R^2 or T^2, and the square root is then taken merely for convenience. Thus z^n and z^{-n} have the same proportional error nt, where t is the proportional error in z.

(*e*) $$Z = z_1 z_2 z_3 \cdots. \tag{14}$$

From Eq. (4) we derive

$$T = (t_1^2 + t_2^2 + t_3^2 + \cdots)^{\frac{1}{2}}. \tag{15}$$

This result applies equally to

$$Z = z_1 z_2 \cdots / z_3 z_4 \cdots, \tag{16}$$

since, as just noted, the proportional error in z_3^{-1} is the same as that in z_3, etc. Hence Eq. (15) applies to all functions, like Eq. (16), in which the independently observed quantities enter only as factors.

(*f*) $$Z = az, \tag{17}$$

where a is a constant. Then

$$T = t. \tag{18}$$

(*g*) $$Z = \ln z. \tag{19}$$

Eq. (4) gives $$R = t; \tag{20}$$

that is, the *absolute* error in $\ln z$ equals the *proportional* error in z.

(*h*) $$Z = \log z = M \ln z = 0.43429 \ln z. \tag{21}$$

Then $$R = Mt; \tag{22}$$

that is, the absolute error in $\log_{10}$ of z is 0.43429 times the proportional error in z.

Equations (21) and (22) may be used to illustrate a situation that frequently arises in practice. Consider any exponential function, such as, for instance, the absorption function (also the function of radioactive decay)

$$y = ae^{-cx}. \tag{23}$$

In checking the validity of this equation, and in evaluating a and c, it is quite customary to plot $\log y$ (to be called y') against x, since a linear plot, with slope $-Mc$ and intercept $\log a$, is thus obtained. In order to get the best values of a and c, suppose that we now make a least-squares solution of the $y' : x$ plot, assuming equal weight for each plotted value of $y'(=\log y)$. This is *not* the desired result, *if* the original observations y are entitled to equal weight. On the contrary, as may easily be proved,[2] the *unweighted* least-squares solution of the $y' : x$ plot corresponds just to the particular *weighted* least-squares solution of the $y : x$ plot for which a weight proportional to $1/y^2$ has been assigned to each observed value of y. *Conversely*, if all the observed y-values are entitled to *equal* weight, and one desires the least-squares values of a and c in Eq. (23), it is necessary to *assign* to each value of $\log y$ a weight proportional to y^2, and then get the least-squares solution of this *weighted* set of values of $\log y$ by the usual equations for a weighted linear curve.

To put this qualitatively—if the average "scatter" of the $y : x$ points does not vary with x so that all points may properly be given equal weight, then the scatter $\delta y'$ of the corresponding $y' : x$ points will be found to *decrease* with increasing y (decreasing x). Hence, one naturally should give special weight (quantitatively, a

[2] The proof is very simple. If each y' is given equal weight, it indicates that we consider its uncertainty $\delta y'$ to be a constant (i.e., independent of the value of y'). Then, by Eq. (22), $\delta y' = M \cdot \delta y / y = constant$, where δy is the uncertainy in y. Hence we have implicitly assumed that δy is proportional to y. But weights should always be assigned inversely proportional to the square of the uncertainty. Hence our assumption of equal weights for each value of y' is an implicit assumption of weights proportional to $1/y^2$ for the values of y. Conversely, if δy should actually be considered as constant, then $\delta y'$ varies as $1/y$ and hence each y'-point should be weighted proportional to y^2.

weight proportional to y^2) to the points in the region of small x (large y) values, in drawing the best straight line through the plotted $y' : x$ data.

Similar considerations apply to many other types of plot. Thus it is, I believe, quite customary in engineering work to reduce the graph of many commonly occurring functions $f(x)$ to a *linear* graph, by plotting a properly chosen function of y against x. The scientific danger in such a procedure is just that previously noted; namely, one is likely (intentionally or unintentionally) to treat the resulting linear plot as though all points were of equal weight, whereas actually they should usually be given quite different weights, the required system of weighting depending on the particular function of y used in the plot.[3] All such systematic weighting represents, of course, a weighting to be applied *in addition* to the appropriate weighting of the original observed values of y.

Let us now consider a function of the directly observed quantities for which the resulting uncertainty is not given so simply. Take the very innocent looking relation

$$1/Z=(1/z_1)+(1/z_2). \qquad (24)$$

The simplest form in which Eq. (4) can here be expressed is

$$R=Z^2[(r_1^2/z_1^4)+(r_2^2/z_2^4)]^{\frac{1}{2}}, \qquad (25)$$

which is more complex than might have been anticipated. The complexity arises from the fact that the explicit expression for $Z[=z_1z_2/(z_1+z_2)]$ shows a *double* occurrence of z_1, one of the measured quantities. Such repeated occurrence of one or more of the measured quantities is often encountered, and in all such cases it is advisable to use the general form of Eq. (4). An attempt to apply any special form, unless the form has been especially derived for the function in question, will almost invariably lead to a false result.

[3] The required system of weighting can always be determined from the appropriate form of Eq. (4), just as is done in footnote 2. As one further very commonly used case, consider $y'=y^n$, where n is often $\frac{1}{2}$ or $\frac{1}{3}$. This is like case (d), and Eq. (13) may here be written $\delta y'/y'=n\delta y/y$. Hence if δy should be considered as constant, then $\delta y'$ varies as y^{n-1} and the y'-points should be weighted proportional to the inverse square, i.e., to $1/y^{2n-2}$. Thus for the parabola $y=ax^2$, which becomes linear by plotting $y'=\sqrt{y}$ against x so that $n=\frac{1}{2}$, the y'-points should be weighted proportional to y (i.e., to x^2).

In connection with Eqs. (24) and (25) it is both instructive and important to note that if we set $z_1=z_2=z$, or $Z=z/2$, and hence $r_1=r_2=r$, Eq. (25) no longer gives the correct result! Thus consider the reduced mass of a diatomic molecule like H_2 or O_2. If z is the mass of either atom, and Z is the reduced mass, then $Z=z/2$. The direct application of Eq. (4) gives $R=r/2$ (the correct result), but substitution of $z_1=z_2=z$ and $r_1=r_2=r$ in Eq. (25) gives $R=\sqrt{2}\cdot r/4$, a result that is too small by a factor of $\sqrt{2}$. In a precisely similar way, if Eq. (5) has the form $Z=z_1+z_2$ and if one assumes $z_1=z_2=z$, so that $Z=2z$, then the correct relation is $R=2r$; but the substitution of $r_1=r_2=r$ in Eq. (6) gives $R=\sqrt{2}\cdot r$, which again is too small by a factor of $\sqrt{2}$. The explanation of this apparent discrepancy is as follows.

If Eq. (5) has the form $Z=z_1+z_2$, then $R=(r_1^2+r_2^2)^{\frac{1}{2}}$. If, by chance, the measurements of z_1 and of z_2 are equally reliable, then $r_1=r_2=r$, and $R=\sqrt{2}\cdot r$. The essential fact is that z_1 and z_2 are here assumed to be *independently* measured quantities, so that in calculating Z, the two errors may add *or* subtract, depending on their relative signs. The factor $\sqrt{2}$ results just from this circumstance. If now, in addition, we find by chance that $z_1=z_2=z$, Eq. (5) must be written as $Z=z+z$, in order to show that we *still* have *two* independently measured quantities, and it is then still true that $R=\sqrt{2}\cdot r$. But if there is only *one* measured quantity z, and $Z=2z$, then any error in z is necessarily *doubled* in Z, and $R=2r$. Hence in this latter case Eq. (6), with $r_1=r_2$, no longer applies. Similarly, in Eq. (24), if there is only *one* measured quantity z, with uncertainty r, the *new* form of the function is $1/Z=2/z$ and Eq. (25), with $r_1=r_2$, no longer applies.

In order to give a numerical illustration of the propagation of errors, let us consider a problem of real present interest. Suppose that we have measured the masses $m_1\pm r_1$ and $m_2\pm r_2$ of two isotopic species, such as Li^7 and Li^6. Suppose that we have measured, in addition, the relative abundance $s\pm r_s$, of Li^7 to Li^6. The problem is to find the corresponding mass (atomic weight) m of the mixture, and its probable error r. A closely related problem is to calculate $s\pm r_s$, on the assumption that the other three quantities have been measured.

The desired functions in these two cases are

$$m=(sm_1+m_2)/(s+1), \tag{26}$$

if m_1, m_2 and s are given, and hence

$$s=(m_2-m)/(m-m_1), \tag{27}$$

if m_1, m_2 and m are given.

By the application of Eq. (4) to the function given by Eq. (26) we obtain

$$r^2=\left(\frac{s}{s+1}\right)^2 r_1^2+\left(\frac{1}{s+1}\right)^2 r_2^2+\frac{(m_1-m_2)^2}{(s+1)^4}r_s^2. \tag{28}$$

Similarly, by the application of Eq. (4) to the function given by Eq. (27), we obtain

$$r_s^2=\frac{(m_2-m)^2}{(m-m_1)^4}r_1^2+\left(\frac{1}{m-m_1}\right)^2 r_2^2 +\frac{(m_1-m_2)^2}{(m-m_1)^4}r^2. \tag{29}$$

Thus, if[4] $m_1=7.01818$, $r_1=18\times10^{-5}$,
$m_2=6.01686$, $r_2=20\times10^{-5}$,
$s=11.60$, $r_s=0.06$,

then, from Eq. (26), the corresponding atomic weight is $m=6.93871$. This is on the physical scale, since the given values of m_1 and m_2 are on that scale. From Eq. (28) the probable error is $r=41.3\times10^{-5}$, and this error arises almost entirely from the assigned probable error r_s in the abundance ratio s, as shown by the fact that the numerical values of the three terms of Eq. (28) are 274.6, 2.5 and 1432.1, each times 10^{-10}.

Let us now assume that the *measured* value of the atomic weight is 6.93871 ± 0.000413, and that the values of $m_1\pm r_1$ and $m_2\pm r_2$ remain as before. Then, from Eqs. (27) and (29) we calculate the corresponding relative abundance as $s=11.60$, and its error as $r_s'=0.0707$. The calculated value of s is necessarily identical with that assumed as the measured value, in the first problem, since Eq. (27) is merely another way of writing Eq. (26). But the resulting calculated error $r_s'=0.0707$ is necessarily *larger* than the original assumed error $r_s=0.06$. This is due to the fact that the errors r_1 and r_2 enter into the calculation of r, and if this calculated value of r is now assumed to be a directly assigned value, they *again* enter into the calculation of r_s', in the second problem of the cycle.

To show explicitly this necessary increase in the error in such a cyclic process, consider the area of a rectangle $c=ab$. If we are given $a\pm r_a$ and $b\pm r_b$, and if t is used, as before, for *proportional* error, then by Eq. (15) we get $t_c=(t_a^2+t_b^2)^{\frac{1}{2}}$. Suppose we now assume that the area $c\pm r_c$ and one side $a\pm r_a$ are the measured values, and we are to calculate the remaining side b and its proportional probable error t_b'. Then $t_b'=(t_c^2+t_a^2)^{\frac{1}{2}}$. If now we use for the assigned uncertainty t_c the value $(t_a^2+t_b^2)^{\frac{1}{2}}$ just calculated, then

$$t_b'=(t_b^2+2t_a^2)^{\frac{1}{2}},$$

which shows explicitly why the *calculated* proportional error t_b' is necessarily larger than the assigned error t_b in the first problem of the cycle.

A fairly common type of mistake in the use of Eq. (4) is its application to functions in which one or more of the quantities appearing explicitly are not independent of the others. Therefore let us now consider certain general cases where quantities of this type appear.

Take, for instance, the straight line

$$y=a+bx, \tag{30}$$

which is made to pass exactly through the two measured points y_1x_1 and y_2x_2, for which the probable errors (of y_1 and y_2) are r_1 and r_2, respectively.[5] The problem is to determine the probable error in y. Now even if we are given the probable errors in a and b, we cannot apply Eq. (4), since a and b are *not* independently measured quantities. On the contrary, it is, in this particular case, y_1 and y_2 that represent such quantities and hence y must first be expressed as an explicit function of y_1 and y_2, before we can apply Eq. (4). This incidentally requires the expression of a and of b as explicit functions of y_1 and y_2, and these latter are just the functions needed to determine the errors in a and b. Hence we first determine these errors, r_a and r_b.

[4] Values of m_1, m_2, r_1 and r_2 from M. S. Livingston and H. A. Bethe, Rev. Mod. Phys. **9**, 245 (1937); see p. 373. Values of s and r_s from A. K. Brewer, Phys. Rev. **47**, 571 (1935).

[5] We assume zero error in the x coordinate. As is well known, the problem becomes far more complex in those cases where the values of both x and y are subject to error.

Since, by assumption, $y_1=a+bx_1$ and $y_2=a+bx_2$, we find

$$a=(y_2x_1-x_2y_1)/(x_1-x_2) \tag{31}$$

and, similarly,

$$b=(y_1-y_2)/(x_1-x_2). \tag{32}$$

The application of Eq. (4) to Eq. (31), where $Z=a$, $z_1=y_1$ and $z_2=y_2$, gives

$$r_a{}^2=\left(\frac{x_2r_1}{x_2-x_1}\right)^2+\left(\frac{x_1r_2}{x_1-x_2}\right)^2. \tag{33}$$

Similarly, by applying Eq. (4) to Eq. (32) we find

$$r_b{}^2=\left(\frac{r_1}{x_1-x_2}\right)^2+\left(\frac{r_2}{x_2-x_1}\right)^2. \tag{34}$$

We now substitute Eqs. (31) and (32) in Eq. (30), and obtain

$$y=y_1\left(\frac{x_2-x}{x_2-x_1}\right)+y_2\left(\frac{x_1-x}{x_1-x_2}\right), \tag{35}$$

which is the desired representation of y as an explicit function of y_1 and y_2. Then with $Z=y$, $z_1=y_1$ and $z_2=y_2$, Eq. (4) gives

$$R^2=\left(\frac{x_2-x}{x_2-x_1}\right)^2r_1{}^2+\left(\frac{x_1-x}{x_1-x_2}\right)^2r_2{}^2. \tag{36}$$

The probable error of the function y naturally varies with x.

A still more complex problem is that of the determination of the error in a, or b, or y, when $y=a+bx$ has been fitted to an entire series of observed y's by the method of least squares. The method of derivation remains, however, precisely the same. To get r_a, the error in a, we first express a as an explicit function of the y's. If all the n values of y are equally reliable and hence are given equal weight, the well-known least-squares value of a is

$$a=(\Sigma y\Sigma x^2-\Sigma xy\Sigma x)/D, \tag{37}$$

where

$$D=n\Sigma x^2-(\Sigma x)^2. \tag{38}$$

There are here n independent values of y and hence Eq. (4), as applied to Eq. (37), has n terms. If r is the assumed probable error of any one of the y's, the resulting probable error of a is found to be

$$r_a=r(\Sigma x^2/D)^{\frac{1}{2}}. \tag{39}$$

Similarly, from

$$b=(n\cdot\Sigma xy-\Sigma x\Sigma y)/D, \tag{40}$$

$$r_b=r(n/D)^{\frac{1}{2}}. \tag{41}$$

Then, by substituting Eqs. (37) and (40) in $y=a+bx$, and applying Eq. (4), one gets

$$r_y=r[\Sigma(x-\epsilon)^2/D]^{\frac{1}{2}} \tag{42}$$

for the error in y at the specified value of $x=\epsilon$. Eq. (42) differs from Eq. (39) only in the substitution of $x-\epsilon$ for[6] x.

As already noted, a common type of mistake found in the literature in connection with the propagation of errors, results from a failure to express the function explicitly in terms of the independently observed quantities. It must be admitted, however, that the independent quantities are sometimes well concealed and it is not surprising that, as a consequence, Eq. (4) is not always correctly applied. An outstanding example of the concealment of independently observed quantities is furnished by the Sackur-Tetrode constant[7] S_0. This is usually written as

$$S_0=R_0'\ \ln\ [(2\pi k)^{3/2}\epsilon^{5/2}/h^3N_0{}^{5/2}], \tag{43}$$

but not one of the quantities on the right-hand side of the equation is independent of the others. In fact, $R_0'=R_0/(J_{15}\times10^7)=\nu_nA_n/(T_0J_{15}\times10^7)$, where ν_n, A_n, T_0 and J_{15} *are* independently measured quantities whose observed values are given in Table *a*, page 59, of reference 7. Similarly $k=R_0/N_0$ and $N_0=Fc/e$, where F, c and e are independently measured. Finally, h must be written as $(h/e)e$, since only the ratio h/e can be measured directly. (ϵ is merely the base of $\ln=2.71828$.) Thus the actual, independently observed quantities are ν_n, A_n, T_0, J_{15}, F, c, e and h/e. But ν_n, A_n and T_0 appear twice, once outside the "ln" and once inside. F and c also appear twice, fortunately both times inside the "ln," while e appears three times, and h and J_{15} appear only once. Obviously the strict application of Eq. (4) to Eq. (43) leads to an extremely complex equation, and I feel fairly confident that

[6] Further material on this matter is given by the writer in Phys. Rev. **40**, 207 (1932); see pp. 224–226.

[7] R. T. Birge, Rev. Mod. Phys. **1**, 1 (1929); see p. 65.

no one had previously even attempted to apply Eq. (4) in this particular case.[8]

[8] As a matter of fact, the published error of S_0 in reference 7 was not correctly calculated. The result should have been given as $S_0 = -11.0533 \pm 0.0036$, in place of the published probable error ± 0.0026 (and $S_0/R_0' = -5.5644 \pm 0.0015$, in place of ± 0.0009). The mistake resulted from the clumsy and very approximate manner in which the relations between h and e were handled in reference 7. A correct treatment of these rather involved relations was first given later [Phys. Rev. **40**, 228 (1932)]. Aside from the fact that h/e, and not h, is to be taken as an independently observed quantity, it may be noted that many of the constants listed in Table *a* (p. 59 of reference 7) are actually more or less interdependent, although this table is presumed to contain only independently measured constants, as contrasted with Table *c*, which contains only derived constants. However, as pointed out at the time, the order of calculation of the constants of Table *a* was so chosen as to *minimize* this interdependence and, as a result, it may fairly be claimed that the application of Eq. (4) to any explicit function of the constants of this table, as was done in compiling Table *c*, will result in no *serious* error in the calculated uncertainty of the function.

Fortunately the preceding example is very exceptional, and in the vast majority of situations encountered in physical science, the correct application of Eq. (4) is a comparatively simple matter. For that reason I believe that all students taking scientific courses, not to mention professional scientists, should be familiar with at least the essential facts of the subject, and should be able to apply them whenever necessary.

Some Aspects of the Electron Theory of Solids

L. A. DuBridge
University of Rochester, Rochester, New York

THE discovery of the wave aspect of the electron may certainly be regarded as one of the greatest achievements in physics of this century. This concept was first proposed by de Broglie, experimentally confirmed by Davisson and Germer and by G. P. Thomson, and later developed by Schroedinger and others into a complete physical theory, on the basis of which nearly all phenomena in the field of atomic and molecular physics may now be treated. To many people, however, *quantum mechanics*, which is the name applied to this theory, still signifies an abstract set of complicated equations with which physicists amuse themselves during their excursions into the very entertaining but probably highly "impractical" realm of subatomic phenomena. Will quantum mechanics ever be of use in "practical" problems?

Without discussing whether or not the field of atomic physics is "practical," we can safely assume that everyone will admit into the category of "practical" studies any investigations of the properties of such things as copper wires, steel rails, tungsten filaments and quartz crystals. Quantum mechanics is now being applied to just such problems. The rise during recent years of a quantum-mechanical theory of crystalline solids (and practically all solids except such things as glass are crystalline) has been a development of major importance in physics, even though it has been somewhat overshadowed by the simultaneous and possibly more spectacular development of nuclear physics.

In view of the fact that most of the external world which we can see and hear and feel is made up of solid bodies, it is rather surprising to realize how little has been heretofore known about the fundamental nature of solids—about the forces that hold them together and about the origin of their mechanical, thermal, electrical and optical properties. An adequate theory of solids would obviously be of great help in understanding what we know experimentally about them and possibly in predicting new properties or in discovering hidden relations between known facts.

It now appears that previous attempts to develop a theory of the solid state were inadequate because the theory of the behavior of electrons was inadequate. When we understand how the electrons in solids behave we will have the clue to many of the properties of solids. Quantum theory is now supplying this clue and we can discuss, in surprisingly simple terms, many of the results to which the theory leads.

This theory differs from the older, so-called "classical" electron theories of solids primarily

in the fact that the electrons, instead of being treated as particles, are now treated as waves. In the classical electron theory of *metals* it was assumed that there were free electrons wandering throughout the metal which behaved collectively as an ordinary gas, obeying the classical statistical laws of Maxwell and Boltzmann. Electron *waves*, however, may behave quite differently. To understand their behavior we need only a general idea of the theory of wave motion and a few simple mathematical relations such as the following:

1. The effective wave-length λ of an electron moving with a momentum p is given by the de Broglie relation

$$\lambda = h/p,$$

where h is the Planck constant. (The same relation holds also for photons.)

2. The kinetic energy E of an electron (or any free particle) is related to its momentum by the relation

$$E = p^2/2m,$$

where m is the mass of the electron (or other particle).

3. The behavior of an electron wave in any region of space is determined by the famous Schroedinger equation which is, in one dimension (the x direction),

$$(\partial^2\psi/\partial x^2) + (8\pi^2 m/h^2)(W - V)\psi = 0,$$

where ψ is the amplitude of the wave, W is the total energy of the electron and V its potential energy. We write this equation merely to exhibit it and to remark that the solutions of it can be found, provided only that we know (besides certain boundary conditions) how the potential energy V of an electron depends upon x. The solutions will represent a wave pattern from which the behavior of the electron can be deduced.

The first attempt to develop a modern theory of solids (which was restricted to metals) was made by Sommerfeld in 1927. The very simple assumption was made that, within a metal block, the potential energy V of an electron was everywhere constant. In the more recent theories of Bloch, Brillouin and others the potential energy is assumed to vary periodically, as it must, in going from one atomic nucleus to the next in the crystal. We will consider briefly the Sommerfeld theory of metals, because it serves as an excellent introduction to the more recent theories, and because it is a first approximation to the exact theory and has been extraordinarily useful and successful in treating such problems as the electron emission from metal surfaces.

The Sommerfeld Theory of Metals[1]

Sommerfeld followed in part the procedure of the classical theory, and assumed that in a metal the valence electrons of the atoms became free to wander at random through the crystal lattice. At the surface of the metal there must be forces preventing their escape. So, if we take the potential energy of an electron outside the metal to be zero, then inside it must have a lower value which is taken as constant and equal, say, to $-W_a$. The variation in V in going across a metal block will thus be as represented in Fig. 1. The potential "wall" at the surface serves to reflect back, into the metal, electron waves which are incident on it from within; thus the metal may be referred to as a "potential box."

The problem of solving the Schroedinger equation for a potential energy variation of this form is similar to the problem of finding the wave patterns for the sound waves that can exist in a vibrating body. Only those wave motions can exist that form standing waves; that is, for which $n\lambda/2 = l$, where λ is the wave-length, n is a whole number and l is the length of the "box." In the 3-dimensional case it is simpler to deal with the wave-number vector $\boldsymbol{\sigma}$, a vector whose direction is the direction of propagation of any wave and whose magnitude is equal to $1/\lambda$. Then, for the x direction, say,

$$\sigma_x = n_x/2l_x$$

and similarly for σ_y and σ_z. But, since $p = h/\lambda = h\sigma$, then $p_x = h\sigma_x = n_x(h/2l_x)$. Also, since the kinetic energy E is equal to $p^2/2m$, we have

$$E = (p_x^2 + p_y^2 + p_z^2)/2m = (h^2/8ml^2)(n_x^2 + n_y^2 + n_z^2),$$

[1] Sommerfeld, Zeits. f. Physik **47**, 1 (1928). For a brief summary of the Sommerfeld theory see Hughes and DuBridge, *Photoelectric Phenomena* (McGraw-Hill), pp. 215–224; Slater and Frank, *Introduction to Theoretical Physics* (McGraw-Hill), Chap. 41. For a fuller account see Brillouin, *Quantenstatistik* (Springer), or the article by Sommerfeld and Bethe in *Handbuch der Physik*, Vol. 24/2.

if we assume a cubical box for which $l_x = l_y = l_z = l$.

Thus, the possible values of the kinetic energy of the electron are quantized; that is, not all values are allowed since the n's must be whole numbers. It turns out that the separation between the allowed values is extremely small, of the order of 10^{-22} electron-volt. Nevertheless, the discreteness of the allowed energy values has an important consequence because of a famous rule, well established in atomic physics, known as the *Pauli exclusion principle*. This states that no two electrons can simultaneously occupy the same *quantum state*. In the present problem a quantum state is specified by the three quantum numbers, n_x, n_y and n_z, and by the spin quantum number. Since the latter may take on only the two values $\pm\frac{1}{2}$ we conclude that each energy level in the metal can be occupied at most by two electrons, provided their spins are reversed.

At very low temperatures the electrons tend to occupy the lowest possible levels. But if there can be only two electrons in each level, and if there are N free electrons in the crystal, then the lower $N/2$ levels will be filled. Since N is a very large number ($\sim 10^{23}$), the energy of the highest filled level will be quite large, of the order of 10 electron-volts, which is 400 times larger than that predicted for room temperature by the classical theory. At higher temperatures some of the electrons are excited to slightly higher energy states, leaving some lower ones vacant.

The distribution in energy of the electrons arranged in this way among the allowed levels is known to be given by the so-called Fermi-Dirac statistics. This is a statistical theory developed to apply to an assembly of particles (a gas) that obey the Pauli exclusion principle. It differs radically from the Maxwell-Boltzmann statistical theory (which applies to ordinary gases) in that it yields quite a different type of distribution in energy (or velocity or momentum) among the particles.

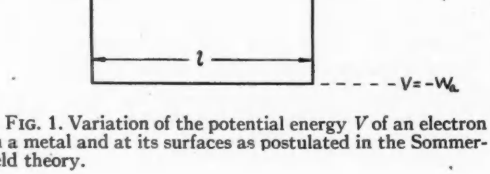


FIG. 1. Variation of the potential energy V of an electron in a metal and at its surfaces as postulated in the Sommerfeld theory.

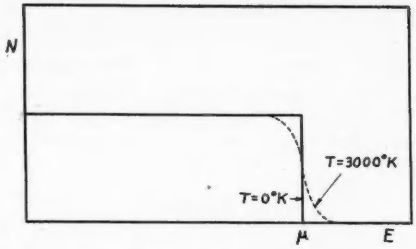


FIG. 2. The Fermi-Dirac distribution function, Eq. (1).

According to these statistics, if we denote by $N(p_x, p_y, p_z)dp_x dp_y dp_z$ the number of electrons per unit volume having momentum components p_x, p_y, p_z within the ranges dp_x, dp_y, dp_z then,

$$N(p_x, p_y, p_z) = \frac{2}{h^3} \frac{1}{\exp\left[(E-\mu)/kT\right]+1}, \quad (1)$$

where E is the kinetic energy and k is the Boltzmann constant; μ is a parameter (determined by the density of the "gas") and is the maximum energy of a particle at 0°K. In Fig. 2, N is plotted as a function of E. The form of the curve shows that at $T = 0$°K all levels are filled up to $E = \mu$ and all higher ones are vacant.

Thermionic emission

We may proceed at once to use Eq. (1) to calculate the number of electrons that escape from the surface of the metal when it is heated. We first point out that the forces at the surface which tend to prevent the escape of the electron must be normal to the surface. Hence, when an electron passes through the surface only the normal component of its momentum is of interest. We may compute the total number of electrons having the momentum component p_x in the range dp_x by integrating Eq. (1) over all values, from $-\infty$ to $+\infty$, of the other components p_y and p_z. We are interested not so much in the number of such electrons per unit volume as in the rate at which they arrive at unit area of the surface. Since this will be proportional to p_x, we will multiply by this factor. If we set $E_n = p_x^2/2m$ and denote by $N(E_n)dE_n$ the number of electrons arriving at unit area of the surface per second with the

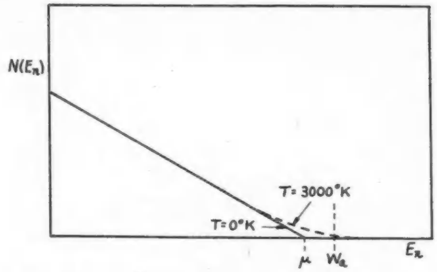


FIG. 3. The "normal energy" distribution function of Eq. (2).

"normal energy" E_n in the range dE_n, we then find[2]

$$N(E_n)=(4\pi mkT/h^3)\times\log\{1+\exp[(\mu-E_n)/kT]\}. \quad (2)$$

In Fig. 3 $N(E_n)$ is plotted as a function of E_n. It is evident that at sufficiently high temperatures the "tail" of this curve will have ordinates of small but appreciable values for energies larger than W_a. Electrons represented by this portion of the curve have enough energy to escape from the surface spontaneously, giving rise to the well-known thermionic emission. The number escaping will be proportional to the area under this curve from W_a to ∞, and when this is calculated we find[2] for the thermionic current,

$$i=AT^2e^{-w/kT}.$$

This is the famous Richardson equation which fits the experimental results with high precision. In this equation A is a universal constant having the value 120 amp cm^{-2} deg^{-2} and $w(\equiv W_a-\mu)$ is called the *work function*. The latter evidently is the energy required to remove from the surface an electron whose initial kinetic energy was just equal to μ. The equation is tested by plotting corresponding values of log (i/T^2) and $1/T$, and this yields a straight line of slope $-w/k$ and $y-$ intercept log A.

Now μ and W_a may both change (very slightly) with temperature so that w may depend on temperature. If we assume that w is equal to $w_0+\alpha T$, the Richardson equation becomes

$$i=AT^2e^{-\alpha/k}e^{-w_0/kT}=A'T^2e^{-w_0/kT}.$$

We will then obtain on a log i/T^2 *vs.* $1/T$ plot a straight line of slope $-w_0/k$ and $y-$ intercept log $(Ae^{-\alpha/k})=\log A'$. The experimental measurements, even for very clean surfaces, do not give the theoretical value of 120 amp cm^{-2} deg^{-2} for the constant A but usually show a value A' of about 60 amp cm^{-2} deg^{-2}. A very small value of α will account for this apparent discrepancy. For contaminated surfaces the values of A' vary widely, often because most of the emission from such surfaces comes from small active regions or is affected by large irregular variations in surface fields.[3]

Photoelectric emission

Suppose now that the metal surface is illuminated by light of frequency ν. *Perfectly* free electrons could absorb energy from the light quanta only by the billiard-ball type of collision, known as the *Compton effect.* In such a collision the energy given to the recoil electron by a photon of visible or ultraviolet light is only a small fraction of the photon energy $h\nu$. This is analogous to the fact that a sphere of large mass (here the electron) can recoil with only a very small kinetic energy when struck by even a rapidly moving sphere of very small mass (the photon); otherwise momentum would not be conserved. However, if the electron is bound to

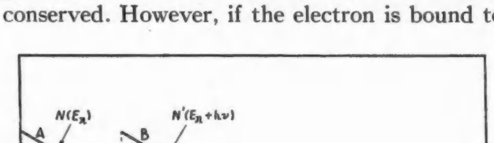


FIG. 4. "Normal energy" distribution curves: A, for normal electrons; B, for electrons that have absorbed the energy $h\nu$.

some third system which can take up momentum (for example, the crystal itself) then the electron may absorb practically the entire energy $h\nu$. Because there is a force field at the metal surface, the electron may be considered as bound to the metal itself, which serves as a third body.

[2] For details see DuBridge, *New Theories of the Photoelectric Effect* (Hermann & Cie, Paris, 1935).

[3] For a more complete discussion of thermionic emission see Becker, Rev. Mod. Phys. **7**, 95 (1935).

Momentum may then be shared between photon, electron and crystal, and the electron may gain the energy $h\nu$. The momentum gained by the electron must, however, be normal to the surface, that is, parallel to the surface forces. The ordinary photoelectric effect is thus made possible by the existence of the surface fields and is therefore called the *surface photoelectric effect.*

Fowler[4] developed an elementary theory of the surface photoelectric effect by assuming that the probability of an electron absorbing the energy $h\nu$ is approximately independent (over small ranges) of the frequency ν and of the energy of the electron. Thus, for the few electrons that gain this additional energy $h\nu$, the normal energy distribution curve would be shifted along the axis by an amount $h\nu$ as shown in Fig. 4.

Let us consider first the situation that exists at 0°K. Evidently, if $\mu+h\nu \gtrless W_a$, some of the electrons will now be able to escape. The quantity $W_a-\mu$ thus equals the minimum value $h\nu_0$ for which a photoelectric current is possible. This quantity is called the *photoelectric work function*, and ν_0 is the *threshold frequency*. The photoelectric and thermionic work functions should thus be equal, and experiment shows that they are equal (neglecting small temperature effects).

For a frequency ν that is larger than ν_0 the number of electrons capable of escaping is proportional to the area of the triangular portion of the curve (shaded in Fig. 4) to the right of W_a. Since the length of the base of this triangle is $h\nu+\mu-W_a=h\nu-(W_a-\mu)=h\nu-h\nu_0$, the number of escaping electrons should be proportional to $(\nu-\nu_0)^2$. Experiment shows this result to be approximately correct for low temperatures and for values of ν close to ν_0.

Also it is evident that, since the maximum energy of the photoelectrons before escape is $h\nu+\mu$, and since in escaping they lose the energy W_a, the maximum energy $E_{\max}$ after escape is $h\nu-(W_a-\mu)=h\nu-h\nu_0$. This is precisely the relation predicted by Einstein and verified by Millikan in his famous photoelectric experiments.[2]

Actual experiments are carried on at temperatures far above absolute zero, whereas the foregoing analysis holds strictly only for 0°K. Fortunately, at room temperature the distribution curves are not very different from those shown and the results remain approximately true. A more exact analysis applicable to any temperature may readily be given, however.[2, 4, 5]

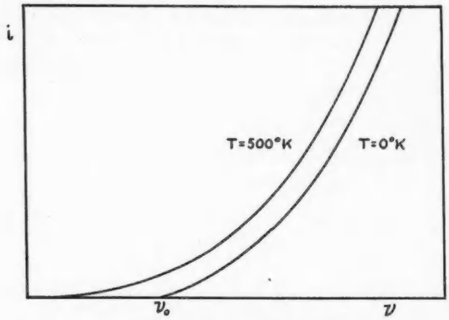


FIG. 5. Spectral distribution curves on the Fowler theory for 0°K and 500°K.

The actual curve for $T \gg 0$°K is also shown in Fig. 4. The "normal" energy distribution of the emitted electrons should have the form of the shaded portion and experiment shows that this, indeed, is its form. However, there is no sharply defined maximum energy of emission at ordinary temperatures; the observed curve merely drops so rapidly that an apparent maximum energy may be deduced by extrapolation. There is, however, a certain "critical energy" E_0, equal to $h\nu-(W_a-\mu)$, which is not equal to the apparent maximum E_m but which can be accurately determined from the shape of the observed curves.[6] Evidently E_0 is given by the Einstein relation while E_m might not be given by it. The difference E_m-E_0 turns out to be approximately constant, however, and so Millikan's method of determining h/e is approximately justified.

The change of photoelectric current with frequency of incident light should also follow a different law at high temperatures as indicated in Fig. 5, and experiments have accurately checked this prediction. The agreement between theory and experiment is strikingly shown in Fig. 6 in which a more convenient method of plotting—a "Fowler plot"—has been used. The experi-

[4] Fowler, Phys. Rev. **38**, 45 (1931). See also reference 2.

[5] DuBridge, Phys. Rev. **43**, 727 (1933).

[6] The quantity E_0 appears as a parameter in the exact equation for the distribution. The value of this parameter may be determined by a graphical method. See references 2 or 5.

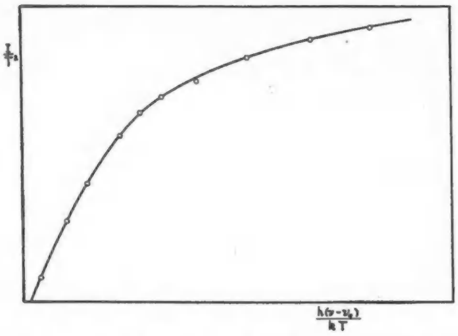


FIG. 6. "Fowler plot" of photoelectric data for a palladium surface.

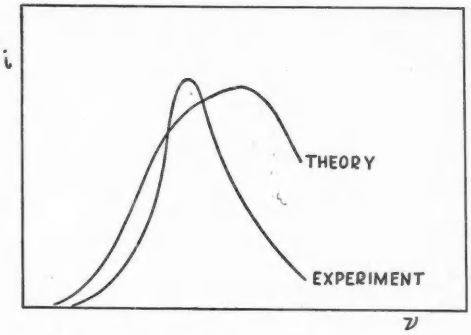


FIG. 7. Theoretical and experimental spectral distribution curves for a sodium surface.

mental points are seen to fall on the theoretical curve with high precision. There is no sharply defined frequency at which the emission begins but there is a characteristic frequency ν_0 (now commonly referred to as the "Fowler threshold frequency") which is given by $h\nu_0 = W_a - \mu$. The value of this frequency ν_0 is determinable from the form of the observed curves[7] and it is not necessary to extrapolate into the axis a curve that approaches it asymptotically, as was always done before this theory was developed.

The foregoing theory, while very useful, is incomplete in that it does not take into account the variation, with energy and frequency, of the probability of an electron absorbing the energy $h\nu$ from the light beam. Calculations of this probability have been carried out on the basis of quantum mechanics[8] and the result is an improved agreement with experiment where frequencies far from the threshold frequency are involved. Over an extended frequency range, experimental curves of photoelectric current *versus* frequency always show a maximum, and the more exact theory predicts such a maximum. A typical set of curves given by A. G. Hill[9] appears in Fig. 7. The agreement between theory and experiment here, however, is not as precise as might be hoped. In addition Hill finds that the theory does not yet predict correctly the form of the energy distribution curves for cases in which $\nu \gg \nu_0$ and in which, therefore, the fastest electrons have considerable energy.

In spite of this incompleteness in the quantitative theory (which leaves interesting problems yet to be solved) the general picture of the emission processes which we now have seems to be fundamentally sound. The present situation is to be contrasted with that in 1925 when no theory was available that could account for the form of the photoelectric distribution curves.

THE ZONE THEORY[10]

It will be recalled that the Sommerfeld theory of metals, discussed in the foregoing section, was based on the assumption that the inside of metal could be considered as a region of constant potential. Actually we know that there must be wide departures from uniformity in potential, particularly in the vicinity of the positively charged nuclei. In addition, this theory *assumed* the existence of free electrons in certain solids—the metals—and gave no explanation of their apparent absence in other solids—insulators or semiconductors. It turns out that when a periodically varying, instead of a constant, potential is assumed within the crystal, as shown

[7] In a Fowler plot (such as Fig. 6) a universal theoretical curve may be drawn. Any experimental curve may be fitted to this by shifting it along the two axes. The amount of horizontal shift required determines the value of ν_0 which is a parameter in the theoretical equation. See references 2 and 4.

[8] Mitchell, Proc. Roy. Soc. **A146**, 442 (1935); **A153**, 513 (1936).

[9] Hill, Phys. Rev. **53**, 184 (1938).

[10] For excellent elementary surveys see Seitz and Johnson, J. App. Phys. **8**, 84–97, 186–199, 246–260 (1937); Shockley, J. App. Phys. **10**, 543 (1939). For more complete treatments see: Slater, Rev. Mod. Phys. **6**, 209 (1934); Shockley, Bell Sys. Tech. J. in press; Fröhlich, *Elektronentheorie der Metalle* (Springer) or Mott and Jones, *Theory and Properties of Metals and Alloys* (Oxford Univ. Press).

in Fig. 8, the behavior of the electrons exhibits new characteristics which at once explain the differences between conductors and nonconductors and which also give the key to the explanation of other solid properties.

When the Schroedinger equation is solved for the case of the periodically varying potential one still finds the discrete, closely spaced energy states as before, given by the condition $n\lambda/2=l$ or (in 3 dimensions) $\sigma_i=n_i(1/2l_i)$ where $i=x$, y or z. However, as we go to higher values of n_i and lower values of λ we come finally to a value of n_i such that $\lambda/2=l/n_i=a_i$ or $\sigma_i=1/2a_i$ where a_i is the distance between atoms in the lattice. For this value of σ standing waves are formed between individual atoms, and such electrons would not be "free" to wander throughout the crystal. Therefore, at this value of σ_i there is a break in the energy level scheme and there is a band of energy values that are "forbidden" to the free electrons. A similar forbidden band exists for each of the values $\sigma_i=m_i/2a_i$, where $m_i=1, 2, 3\cdots$. The allowed energy values are thus divided into bands or zones with relatively large forbidden regions between, as illustrated in Fig. 8. This is in sharp contrast to the continuous range of allowed values in the Sommerfeld theory. (One should not confuse the extremely small "forbidden regions," 10^{-22} electron-volt, between individual *levels* with the large forbidden regions, of the order of 1 electron-volt, between the *bands*.)

The fact that certain energy ranges are forbidden to the electrons in a crystal seems a bit mysterious at first, but finds a ready analogy in the well-known behavior of x-rays. We know that if x-rays of varying wave-length are allowed to fall on a crystal in a given direction (say, normal to the crystal planes), certain wave-lengths are totally reflected (Bragg reflection). X-rays of this wave-length are thus "forbidden" to enter the crystal, because of interference. The Bragg reflection condition is $m\lambda=2a\sin\theta$. Hence, for $\theta=90°$, $\sigma_x=m/2a_x$, where a_x is the distance, in the x direction, between crystal planes. This is precisely the situation for the electrons; if the wave number for an electron satisfies the foregoing relation it cannot be propagated through the crystal.

Since the individual levels are determined by the condition $\sigma_i=n_i/2l_i$, while the first forbidden zone ($m=1$) falls at $\sigma_i=1/2a_i$, it is evident, as we have seen, that at this break n_i has the value $N=l_i/a_i$. Hence, there are just N allowed levels in each band for the 1-dimensional case and N^3 possible levels for a 3-dimensional cubic crystal. But N is just the number of atoms along each edge of the crystal and N^3 is the total number of atoms, N_a, in the crystal. *Each allowed zone, therefore, contains a number of levels equal to the number of atoms in the crystal.* Since each level may be occupied by not more than 2 electrons, a band is filled if it contains $2N_a$ electrons.

In addition, it can be shown that there is a correspondence between each allowed zone of levels in the crystal and one of the sharply defined energy states of the isolated atoms. One can say, then, that the sharply defined energy levels of an atom broaden out into bands when the atoms are coalesced into a crystal. The broadening is great for the valence electrons, and very slight for the inner electrons as shown in Fig. 8.

Electric conductivity

This zone structure at once leads to the answer to the question of why some crystals are insulators and some conductors. Suppose, first, that each atom contains an *even* number of electrons. Then a group of the lower zones will be completely filled, each with $2N_a$ electrons, and the upper ones completely vacant. But if there is a large forbidden band between the top of the highest filled zone and the bottom of the next empty zone [Fig. 9(*c*)], then an electric field of ordinary strength cannot give the electrons enough energy to jump this gap. In other words, electrons cannot be accelerated by the field, and the crystal is an insulator. Thus the

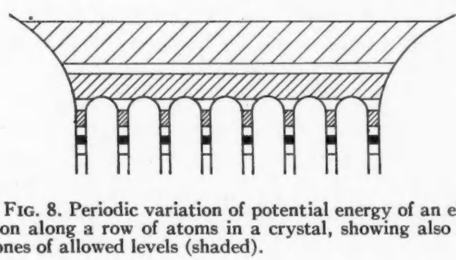


FIG. 8. Periodic variation of potential energy of an electron along a row of atoms in a crystal, showing also the zones of allowed levels (shaded).

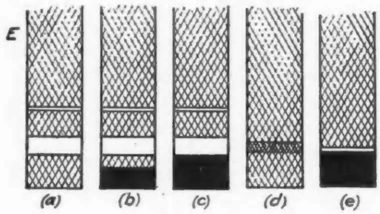


FIG. 9. Possible energy spectra for valence electrons in crystals. With allowed zones shown as shaded in (*a*), the crystal will be a conductor if the lower band is half-filled as in (*b*), or an insulator if completely filled as in (*c*). If the zones overlap as in (*d*) the crystal will always be conducting. If the forbidden region is small as in (*e*) it will be a semiconductor.

electrons in any filled zone, while they are free to wander at random through the crystal, are not "free" in the sense that they can take part in electric conduction. They cannot be accelerated and thereby be shifted to higher energy states, since for *every* electron the states immediately above are either already occupied or are forbidden.

If, on the other hand, the crystal has an *odd* number of electrons per atom, so that the uppermost occupied band is only half-filled [Fig. 9(*b*)], the electrons in this band *may* move to higher energy states and hence are "free" to take part in electric conduction. Such a solid will be a metal and the electrons in it will behave very closely according to the Sommerfeld theory, which therefore becomes a special case of the more complete theory.

This would lead one to expect that all crystals made up of atoms of odd atomic number will be conductors—which is probably true unless "molecular crystals" are formed—but that crystals composed of atoms of even atomic number will be insulators, which is certainly not true. Not all even-element crystals are insulators because in some cases the allowed bands of levels overlap [Fig. 9(*d*)], thus obliterating the forbidden regions and allowing conduction. This is the situation for all the divalent metals.

If the forbidden region between the uppermost filled zone and the next empty one is very small [Fig. 9(*e*)], transitions across the region may take place and the crystal will be a semiconductor, that is, one whose conductivity is small but increases rapidly with temperature.

Optical properties

A solid that contained an atmosphere of *perfectly free* electrons (such as postulated in the Sommerfeld theory) would not absorb light but would totally reflect it from its surface. This is because, as we have seen, momentum must be conserved in a collision between photon and electron. Most metals behave in this way for ordinary light, but probably all metals and insulators show marked selective absorption of light in certain regions of the spectrum. This can be understood at once on the zone theory of solids. In order that an electron may absorb light of a given frequency it is necessary that there be some mechanism by which the electron may transfer momentum to the crystal. Such a mechanism is available at the surface through the field existing there, and we have seen that this makes the surface photoelectric effect possible. Within the volume of a crystal, however, another mechanism is required, and this is found in the periodic force field which has been postulated.

There is, however, a "selection rule" which governs the possible momentum transfers to the crystal lattice. This rule may be understood by referring again to our analogy between the behavior of electron waves and x-rays. When a beam of x-rays strikes a crystal face normal to a set of crystal planes it will be totally reflected if the Bragg condition is satisfied, that is, if $n\lambda/2=a$ or $\sigma=n/2a$. Since the momentum of an incident photon is $h\nu/c$, the momentum transferred to the crystal on total reflection is $2h\nu/c=2h\sigma$. Using the foregoing relation between σ and a we get for the momentum transfer $\Delta M=2hn/2a=n(h/a)$. The allowed momentum changes are thus integral multiples of the quantity h/a. In a similar manner an electron may gain energy from a photon in the crystal lattice, provided its momentum change in the process is a whole multiple of h/a. The change in wave number σ of the electron must then be a whole multiple of $1/a$. This quantity is just twice the width of a zone. Hence an electron with a wave number σ in the lowest zone (whose upper limit m_i is 1) must, after absorbing a photon, have the wave number $\sigma'=\sigma\pm(1/a)$. This will take it either into the second zone

$[|\sigma'|=(1/a)-\sigma=2(1/2a)-\sigma]$ or the third zone $[|\sigma'|=2(1/2a)+\sigma]$. Transitions from one level to another in the same zone will always be forbidden.

The corresponding allowed changes in the energy of the electron will depend upon the relation between energy and wave number. For an electron in free space the relation is $E=p^2/2m=(h^2/2m)\sigma^2$. This relation is plotted in Fig. 10(a). Within the crystal, however, this relation is radically altered, showing discontinuities at values of σ given by $\sigma=\pm m/2a$. The relation between E and σ will then be of the form[11] shown in Fig. 10(b). By translating the various segments of this curve to the right or to the left by the amount $\pm 1/a$, we obtain the "reduced zone plot" shown in Fig. 10(c). On this reduced plot the allowed electron transitions induced by light absorption are represented by vertical lines such as those indicated by $h\nu_A$, $h\nu_B$ and $h\nu_C$.

The zone scheme for a crystal then correlates at once with its optical properties. Thus, if the lower zone should be only partially occupied by electrons up to the point A, then the light of lowest frequency which could be absorbed would correspond to the transition $h\nu_A$. The largest energy change from first to second zone would be $h\nu_B$, so the crystal would have an optical absorption band extending from ν_A to ν_B. The frequency ν_A will commonly be in the ultraviolet. However, for an insulating crystal for which the first zone is filled, the smallest energy change will be $h\nu_D$, which may be in the visible part of the spectrum, and the absorption band will extend from ν_D to ν_B. Optical absorption measurements will thus give direct experimental evidence for the form of the zone patterns.

Volume photoelectric effects

An electron in a crystal which has undergone an allowed transition will be transferred to an energy level that lies in an unfilled portion of a zone, transitions to levels already occupied being, of course, impossible. These electrons will be essentially conduction electrons; hence, when an insulating crystal absorbs light, it will be rendered slightly electrical conducting. This is the phenomenon of *photoconductivity*.

In the case of *metals* the additional conductivity induced by light will not be noticed. However, in metals these excited electrons may, under proper conditions, reach the surface and escape, contributing appreciably to the photoelectric emission from the surface. Thus we see that the photoelectric current emitted by a

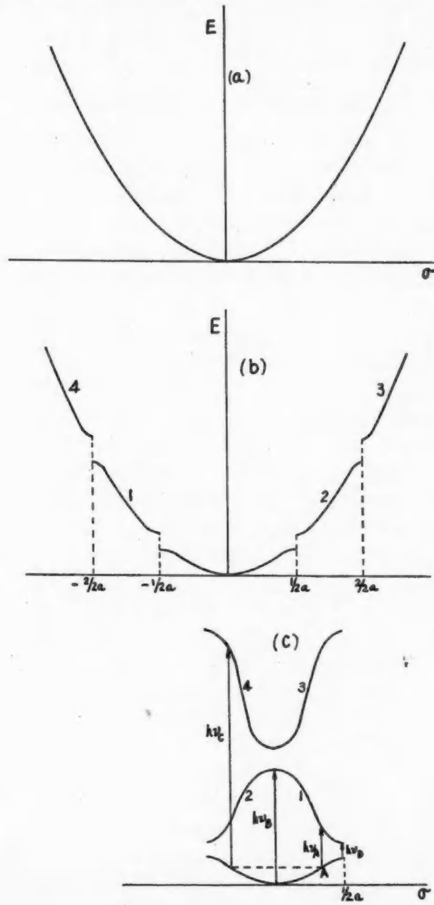


FIG. 10. Energy *vs.* wave number diagrams: (a) for free electrons; (b) for a crystal, showing three zones; (c) "reduced zone plot," showing allowed optical transitions. Segments 1, 2, 3 and 4 of the curves in (c) correspond to similar numbers in (b).

[11] The exact form of this curve depends greatly on the structure of the crystal. It will be noted, however, that though there may be large gaps in the values of E, there are no gaps in the values of σ. Electrons having σ-values near the discontinuities do not behave at all like free particles, and may behave as though they had an infinite or even a negative (!) mass. There is no familiar analogy for such a behavior any more than there is for many other quantum-mechanical phenomena. It is connected, however, with the impossibility of forcing the electron to a forbidden energy level.

metal surface may be composed of two components: (1) a so-called "surface effect," caused by light absorption by electrons in the surface field, for which selection rules do not apply; (2) the "volume effect," caused by allowed transitions between zones within the crystal. The "threshold frequency" for the volume effect will correspond to ν_A of Fig. 10(*c*)—provided the upper state so reached is above the surface potential barrier—and normally will be much higher than the surface threshold frequency ν_0. Unfortunately, few photoelectric experiments have been carried out sufficiently far in the ultraviolet to secure good measurements of the contribution of the volume effect. This is an interesting subject for future investigations.

Other crystal properties; some difficulties

Space will not permit discussion of other properties of crystals—thermal, thermoelectric, mechanical, ferromagnetic, fluorescent, etc.—on which the new theories have shed much light. It is hoped that many readers will consult the literature[10] on the subject for further information. One can scarcely help being impressed with the great power of a theory which has accomplished so much in so many directions in such a short time.

It would not be fair to assert that all problems are solved—actually, only a beginning has been made on most of them—or that there are no difficulties ahead. The difficulties apparently are not of principle but are no less serious. Like all difficulties in physics, they are of two types: mathematical and physical. The mathematical difficulties of calculating the electric fields in a crystal and the corresponding zone patterns are really impressive. Exact calculations, even for the simplest crystals, are extraordinarily tedious. For more complex crystals, for alloys and mixed crystals, exact calculations are at present out of the question; but qualitative information can be obtained which may be of great usefulness.

On the physical side one faces at once the fact that no actual crystal has the perfectly and uniformly periodic structure assumed in the theory. There are small and large breaks in the crystal pattern; there are minute traces of foreign materials which spoil the perfect regularity. These irregularities will alter the zone pattern for the electrons. For many crystal properties these minor alterations are unimportant. But other properties are extremely sensitive to just such irregularities. Nevertheless, much has been learned even about these properties by comparing the actual behavior of crystals with that expected for a perfect crystal. One can interpret, in some cases quite satisfactorily, the role played by irregularities and impurities in the lattice.

This discussion of solid state theory necessarily leaves much unsaid. Its purpose has been only to afford a glimpse into a realm of fascinating interest, with the hope that some readers will be attracted to further exploration. In doing this a certain amount of oversimplification has been necessary. But the student entering this field should see that the fundamental physical ideas involved *are* simple and, at the same time, powerful. They are so simple and so important that there is no longer any reason why undergraduate students, for example, should not learn to think of solids in terms of the new theory. One may as well avoid the later necessity for "unlearning" a discarded theory which is no simpler than the new theory and which leads to wrong results.

My father took me sometimes to see masons, coopers, braziers, joiners, and other mechanics, employed at their work, in order to discover the bent of my inclination, and fix it if he could upon some occupation that might retain me on shore. I have since, in consequence of these visits, derived no small pleasure from seeing skilful workmen handle their tools; and it has proved of considerable benefit to have acquired thereby sufficient knowledge to be able to make little things for myself, when I have had no mechanic at hand, and to construct small machines for my experiments, while the idea I had conceived has been fresh and strongly impressed on my imagination.—Autobiography of Benjamin Franklin.

The Operational Theory in Elementary Physics

V. F. LENZEN
Department of Physics, University of California, Berkeley, California

I

IN recent years logical criticism of physics has engaged the interest of physicists. Programs have been undertaken to analyze the procedures of physics and to achieve an adequate understanding of the meaning of physical concepts. A goal for this critical activity was set by Mach, who banned rationalistic metaphysics from physics and sought to restrict physical description to the data of perception. The first definite achievement of criticism in physics was represented by the relativistic concepts of space and time of the restricted theory of relativity; the widespread interest of physicists in critical activity was facilitated by Bridgman's operational theory of physical concepts; and the latest achievement is the interpretation of the concepts and laws of quantum mechanics as instruments for the prediction of the results of experiments. New physical discoveries find their way into the textbooks. Should the critical mode of interpreting physics be expressed in the elementary treatments of the subject? In this paper I propose to consider to what extent the operational point of view should be used in elementary physics.

Let us first consider a statement of the operational point of view. According to Bridgman, classical physical concepts were assumed to express properties, whereas the operational point of view assumes that they express operations. The definition of a physical concept is expressed by a description of the operations by which its applicability to phenomena is ascertained. Adopting a quantitative formulation, the definition of a physical quantity is the description of the operations by which it is measured. The interpretation of quantum mechanics has presupposed the point of view that the aim of physical theory is the prediction of the results of experiments. The concepts of theoretical physics are designed to facilitate the control of experiments and the interpretation of their results. Such, then, is the operational theory of physical concepts. Since this point of view now appears to be generally accepted in physics, it is appropriate to ask if any modification is required in the exposition of elementary physics.

As a preliminary summary I should say that criticism of elementary physics from the operational point of view reveals that in traditional expositions the disposition is to *treat physics as mathematics, to define concepts in terms of metaphysics instead of operations, to fail to recognize the element of convention in physical theory, and to define units for the measurement of physical quantities prior to an account of the experiments in which they are applied.*

II

A first subject for revision in the light of operational theory is the relation of geometry to physics, more especially, the relation of geometry to mechanics. The traditional point of view is expressed in the following quotation from Slate's *Principles of Mechanics* (1905). He says,

> In the first two chapters we shall be occupied with conceptions—Velocity and Acceleration—that rest entirely upon a mathematical basis. . . . If mechanics is taken to include kinematics also, as it frequently is, that part of the science which is physical and not geometrical must be specially distinguished. It is designated as Dynamics. The point should be watched at which the transition is . . . made by introducing experimental results in the framework of our science.

The ideas expressed by Slate are characteristic of older expositions of mechanics. In the study of motion there is recognized the progression: geometry, the science of space; kinematics, the science of motion which is based upon the addition of time to space; dynamics or mechanics, which explains the motions of the material bodies in the physical world. Geometry and kinematics are viewed as mathematical sciences; dynamics or mechanics, as a physical science. Since mathematics traditionally has been viewed as a rational science that is independent of experience, these older works imply that there is a sharp distinction between geometry and mechanics.

The interpretation of geometry as a non-empirical science may be traced to two important

philosophies of geometry. The Platonic theory interpreted geometrical figures to be ideal forms that have supersensuous reality, and the Kantian theory taught that geometrical figures are constructed in pure intuition, independently of sensory experience. In opposition to the non-empirical point of view is the theory that, insofar as it has any relevance to physics, the subject matter of geometry is the spatial properties of configurations of physical things; the metrical structure of space is defined by the possibilities of relative position of practically rigid bodies. Accordingly, geometry is viewed as an experimental science, indeed, as the most general branch of physics.

The preceding discussion provides a basis for the definition of the concepts of geometry in terms of operations. Consider the concept of length. A qualitative concept of length expresses an intuitive property of a body that is experienced in perception. This furnishes the basis for the classical physical concept of length as an intrinsic property of a body which is absolute and independent of relations to other bodies. To be sure, the operational point of view must permit us to recognize that bodies have length as an intuitive character. However, the physical concept of length is precisely defined only by a description of the operations of testing equality in length and of measuring lengths. In view of the relativistic theory that length is relative to a frame of reference, the discussion of length in geometry should begin by postulating that all bodies are at rest in a selected frame of reference. Two straight rods are equal in length if the end points of one can be brought into coincidence with the corresponding end points of the other. The length of a rod relative to a standard is determined by the operation of laying off the standard of length on the rod and counting the nearest number of times that the standard will go into the rod. The preferred definition of physical length in terms of these operations is that "the length" of the rod is the numerical measure relative to the standard. The operational point of view appears to interpret a physical quantity as a number which is assigned to an object by certain operations. This view was set forth by Eddington, as well as by Bridgman. Eddington says,[1] "Consider, for example, a length or distance between two points. It is a numerical quantity associated with the two points; and we all know the procedure followed in practice in assigning this numerical quantity to two points in nature." Again, he says, "A physical quantity is defined by the series of operations and calculations of which it is the result."

To summarize: I think that an essential element in the teaching of mechanics is the exposition of the role of geometry as the fundamental branch of physics; it is the theory of the positional relations of practically rigid bodies. Cartesian coordinate systems are to be visualized as cubical lattices formed out of rigid rods that are equal in length. Such structures provide frames for the location of events in the physical world. Descriptions of the operations of testing equality in length and of measuring lengths furnish the simplest examples of the operational definition of physical concepts. Since the metrical structure of space is defined by the positional relations of practically rigid bodies, the student should not be astonished at the hypothesis that in some regions physical space is Euclidean to the first approximation only. In this way the mind may be prepared for a later understanding of relativity.

III

The critical analysis of physics has exhibited the role of operations; it has also demonstrated the fundamental significance of definitions which are posited by convention. Accordingly, it is important to understand the function of definitions in operational procedure. Since this problem does not appear to have been generally studied, I shall have to rely primarily upon my own views.

The role of definitions in interpreting operations may be exemplified by mechanics. Archimedes created statics as a deductive system, adopting as a fundamental assumption the *restricted principle of the lever*. This principle states that if equal weights are attached to the ends of a lever with equal arms, the lever in a horizontal position will be in equilibrium. From this and other assumptions Archimedes derived a principle of the lever for unequal arms. What is the significance of Archimedes' restricted principle? He assumed that the principle is self-evident, but

[1] *The Mathematical Theory of Relativity* (1923), p. 1.

Mach later contended that it is a generalization from observation. In order to test whether or not the principle is an experimental result, let us examine how one could verify it. For this discussion I shall assume that we know how to measure length. In order to construct a lever, one would choose a straight rigid rod of uniform cross section and place it on a fulcrum so that both lever arms are equal in length. One would then hang equal weights from the two ends of the arms and observe that the lever remains horizontal, thereby verifying the restricted principle of the lever experimentally. But now I should like to ask the experimenter how he knows that the two bodies attached to the ends of the lever are equal in weight. He could reply that both bodies were stamped with the same number, for example, *100*, indicating the mass in grams. Since weight is proportional to mass, the two bodies are equal in weight. But how does the experimenter know that the standard weights have been accurately made? He may reply that the weights were obtained from a reliable instrument-maker. Granted that the maker was honest and accurate, how does he know that he should stamp the same number on both bodies? I think that if one observed the instrument-maker in his workshop one would discover that he employed a beam balance in order to test the equality in weight of the two bodies, that he tested the equality of each to a standard body of 100 gr wt. In other words, the instrument-maker probably employed the restricted principle of the lever as a definition of equality of weights, and realized this definition in the system consisting of the balance and the two bodies in equilibrium. If, as a next step, he declares that the two bodies are equal in mass, he employs as his definition of equality in mass, its proportionality to weight. In statics we use the concept of weight, so that the principle of the lever is a definition of equality in weight.

The foregoing discussion illustrates how an operation must be interpreted by a convention in order to yield the definition of a physical quantity. It must be noted, however, that the definitions which are posited by convention, are founded on the observable behavior of the apparatus. Accordingly, our discussion should be completed by describing the empirical basis of the principle of the lever. Since the principle has been inherited from antiquity, it requires careful analysis to discover the basic generalization from experience. The empirical law is that if two bodies maintain one equal-arm lever in equilibrium, they will maintain all others in equilibrium, regardless of position, length, material, color and so forth. This empirical law is transformed into a principle which defines equality of weight in statics. In other words, the self-evidence which Archimedes attributed to his principle is a consequence of its status as a definition in the statics that he founded. In general, the quality of self-evidence is likely to indicate a definition.

The preceding discussion demonstrates that definitions in the form of principles are necessary to interpret the results of operations. The operation is a means of creating conditions under which it is possible to test whether or not a concept is applicable to the physical world. From the point of view of the present discussion, if a student tests the restricted principle of the lever by using two equal weights from a set of standard weights, he is in reality testing the equality of the standards; the principle of the lever is not verified but is used as a definition. Of course, one could use another definition of equality in weight, and then the principle of the lever would become an experimental law. But the alternative definition would be in terms of some other principle of mechanics serving as a definition and its application would require a different operation. For example, we may define two bodies as equal in weight if they stretch a specific spring the same distance. In this case our definition of equality in weight would be based upon the law that the force is proportional to the stretch. This discussion of the role of conventions in operations indicates that we should note whether an experimental operation is intended to prepare the exemplification of a definition or to obtain an experimental test of some law or principle.

IV

The operational methods employed in statics may be applied to the fundamental concepts of dynamics. In old fashioned textbooks mass is defined as the quantity of matter in a body, and force is defined as the cause of acceleration. From the critical point of view such definitions are metaphysical. Extremists would declare them

to be verbal. We must describe operations by which we determine whether or not these concepts are applicable. Quantitative definitions must describe operations by which mass and force are measured. There are, of course, many modes of procedure.

Let us begin with the definition of mass. Mach made an important contribution to this problem by conceiving an experiment in which two isolated bodies interact. The respective accelerations are measured, and their ratio is found to be constant; that is, $a_1/a_2=k$. One may now define the ratio of the masses of the two bodies by the equation $m_2/m_1=-a_1/a_2$. If we decide arbitrarily that the body with mass m_1 has unit mass, the numerical value m_2 is determined for the other mass. Thus, assuming that we have agreed upon the procedure for measuring acceleration, we can determine the mass of a body relative to a unit by a calculation from measures of accelerations during an interaction.

The procedure for the comparison of masses by interaction is suitable for demonstrating that the concepts of physical properties are based upon the actions and reactions of bodies in experiments. Some years ago my colleague, Professor W. H. Williams, used the word "behavioristic" to describe the fact that the properties of physical systems are determined from their behavior, that is, from their responses to stimuli. This psychological analogy expresses the fact that the properties assigned to bodies are modes of reaction under specific physical conditions. Weyl[2] has formulated this point of view in the principle: "That we regard the result of the measurement read from a reaction as a property pertaining to the body under observation in itself, if the result of the measurement does not change upon change of the conditions of the reaction; the assumption being that this body every time enters into it in the same state." Prior to the understanding of the operational point of view, the disposition was to attribute physical properties to systems absolutely, and not relatively to experimental situations. The definition of mass as quantity of matter exemplifies the absolutistic point of view. Since matter was defined as material substance it was difficult to understand how mass could be relative according to the restricted theory of relativity. If the operational point of view is impressed upon students in their formative period, it should then be easy for them to grasp Bohr's principle of complementarity. In atomic physics, physical properties exhibit themselves only through the mediums of interventions and reactions which are interpreted in terms of laws that are postulated to be valid. Complementarity is based on the fact that the interaction which serves to measure one property makes the measurement of a second property impossible in principle.

Resuming the discussion of the problem of the concepts of dynamics, it is clear that, having defined mass in terms of the experimental result that the ratio of accelerations is constant, we may define force by the equation $F=ma$. From the definitions of mass and force Newton's third law of motion follows immediately. On reviewing the definition of mass one observes that the empirical foundation for the definition of mass is the empirical basis for the third law of motion. In effect, the third law of motion furnishes a definition of the ratio of masses.

An alternative operational procedure in dynamics starts with the definition of momentum through the principle of the conservation of momentum. A moving body may be defined to have momentum. Two bodies have momentums that are equal in magnitude but opposite in direction, if on collision they adhere and are both brought to rest. A moving body which is similarly brought to rest on collision with two bodies having equal momentums is defined to have a momentum twice the magnitude of each of these two bodies but opposite in direction. Having defined momentum in terms of collision experiments and the principle of conservation, which indeed is only another method of using the third law, we may then define force as the time-rate of change of momentum. Mass may be defined through the empirical law that for speeds small in comparison to the speed of light the momentum of a body is directly proportional to its velocity. The mass of a body may be defined as the factor of proportionality. The advantage of the present procedure is that it may be adopted for the restricted theory of relativity. Moreover, the student has nothing to unlearn if he begins with time-rate of change of momentum instead of ma.

[2] *Mind and Nature* (1934), p. 3.

Instead of beginning with the concepts of mass or momentum, one may start with force. A good procedure is to adopt, as a typical force, the force exerted by a stretched spring. The definition of the measure of force is the law that the force exerted is proportional to the stretch. The second law of motion may then be used as a definition of mass or momentum.

In critical discussions of dynamics one observes different interpretations of the second law. In one context it serves as a definition of force, in another it serves as a definition of mass or momentum. Its logical status fluctuates. A unified interpretation of the second law may be obtained by an alternating procedure in the employment of the law as a definition. We may assume that force is exerted by a stretched spring and agree to use the relation between force and stretch as a definition of the measure of force. We thereby restrict the discussion to a particular type of force. If such a force is applied to a particular solid body, the body is accelerated, and it has been found that, for the speeds considered in classical mechanics, there is a direct proportionality between the force and the acceleration. The results of experiment may be expressed by $F=ma$, where m is a factor of proportionality which depends on the particular body accelerated and is called the *mass* of the body. The second law is then exhibited as a generalization from experience which has been transformed into a definition of an intrinsic property of a body, its mass. Assuming that the same force is realizable at different times and places, we find the mass of a body to be independent of time and position.

Let us now assume that we have attributed invariable masses to a set of bodies. We shall subject a body to various physical conditions so that it is accelerated. It may be allowed to fall freely, it may be pushed by another body, it may be given an electric charge and placed in the vicinity of another charged body. We may explain the acceleration as a consequence of the action of force which is defined by $F=ma$. The second law now serves as a more general definition of force than the force of a spring which was used to define mass. The empirical generalization which provides the basis for the employment of the second law as a definition of force is that similar physical conditions have been found to be correlated with the same product of mass and acceleration, regardless of the body accelerated. It is a further fact that simple laws of force can be found for force as now defined. The law of gravitation is probably the most notable example of a law that expresses force as a function of the masses and relative positions of the bodies involved. The interpretation of the second law as a definition of force is especially exemplified by the introduction of frictional, electric and magnetic forces.

The foregoing procedure may be summarized as follows. A restricted concept of force was assumed, and the second law was employed to define the concept of mass. The new concept of mass was used with the second law to define a more general concept of force. By an alternating procedure the second law was used to define first one concept and then the other. The procedure appears to be circular, but the same concept may be applied with different degrees of generality. In this process, which I have called *successive definition*, physical concepts may be defined initially by special laws and, as new generalizations are attained, the new laws may be transformed into definitions of more general concepts. As I have shown elsewhere,[3] the process of successive definition may be illustrated by ordinary algebra, geometry, statics, dynamics, electrodynamics and thermodynamics.

It may be contended that the foregoing discussion of dynamics is in terms of ideal experiments, of thought-experiments which no one ever performs. These experiments would serve to build up dynamics in imagination by a fictitious operationalism. It may be questioned whether such an imaginary reconstruction of a science is suitable for the exposition of elementary physics. It would appear that in teaching it is preferable to use an operational procedure that essentially reproduces the historical process or present practice. However, the experience gained with the ideal experiments should assist us in working out a critical procedure which conforms to present practice.

I suggest that we begin with weight as the typical force and define it in terms of operations with the lever and the restricted principle of the lever. We may then define mass by the law that it

[3] *The Nature of Physical Theory* (J. Wiley, 1931).

is proportional to weight, but must add the postulate that mass is invariant in a displacement. Having assigned numerical measures to the masses of bodies, we may then use the second law of motion to define force. Whenever a mass is accelerated we say that a force is acting which is defined by $F=ma$. It may be remarked that this definition involves the hypothesis that it is possible to express force as a simple function of position or other properties. Elastic force that obeys Hooke's law and the force of gravitation are examples. That classical dynamics is valid for macrophysical phenomena means that the laws of motion, which are definitions based upon the empirical laws involved in operations, are suitable for a relatively simple description of phenomena.

V

I turn now to another field of physics, that of heat. I think that operational analysis is especially needed in this field. The usual order of exposition is that units of heat and specific heat are defined and then the method of mixture is described as a phase of calorimetry. It appears to me that the operational point of view demands that the description of the method of mixtures precede the definitions. Definitions are founded on experiments.

The critical discussion may be initiated by considering the status of the fundamental principle of calorimetry, that during the exchange of heat between two bodies or systems the heat lost by one system is equal to the heat gained by the other. How do we know that this principle of the conservation of heat is true? At first sight one is inclined to say that it is an experimental law which can be confirmed by the measurements made during a calorimetric experiment. In order to calculate the quantities of heat lost or gained in such an experiment it is necessary to know the specific heats of the substances involved. These may be found from the tables, but how were the tables obtained? It is probable that the specific heats were determined by the principle of conservation of heat from data furnished by calorimetric experiments. The principle that heat lost is equal to heat gained thus provides a definition of specific heat in terms of the operations with a calorimeter.

A possible mode of procedure is the following. Consider two bodies of different materials that have masses m_1 and m_2. Let these bodies be placed in contact and the system insulated; experience shows that they will come to temperature equilibrium. Let Δt_1 be the gain in temperature of m_1 and Δt_2 the loss in temperature of m_2. If during thermal interaction there is no change of state, experiment shows that $m_1\Delta t_1/m_2\Delta t_2=\text{constant}$. We may write this constant as s_2/s_1, and obtain

$$m_1\Delta t_1/m_2\Delta t_2=s_2/s_1.$$

Then s_1 is characteristic of the substance of which *1* is composed, and s_2 is characteristic of *2*. For an arbitrarily chosen substance s may be fixed as unity and the values of s for all other substances may be calculated. The symbol s designates the specific heat of a substance. The concept of specific heat expresses the mode of temperature reaction of a substance to some other substance with which it is in contact and is defined by an empirical law obtained from a calorimetric experiment. We may derive

$$m_1s_1\Delta t_1=m_2s_2\Delta t_2.$$

If we define increment of heat as $ms\Delta t$, we obtain the principle that *heat lost = heat gained*. We have defined our concepts of specific heat and heat in the light of experience with operations so that experiment confirms a principle of conservation.

Conclusion

I have offered examples of procedures for constructing definitions in mechanics and the theory of heat. In every case an empirical generalization from the results of experiments furnishes the basis for a definition. It does not seem to me that this point of view is adequately utilized in our textbooks. The creators of physical concepts *did* employ the operational procedure. Newton explicitly stated that the properties of bodies are to be derived from experiments. Thus physical concepts have been introduced into physics through operations. But once the science had been developed, a deductive mode of exposition proved to be attractive. Metaphysical interpretations also were introduced. The operational point of view has the appearance of novelty only because many physicists seem to have forgotten the historical modes of formation of habitually employed concepts.

The Laws of Electromagnetic Induction

W. V. Houston
California Institute of Technology, Pasadena, California

THE subject of electromagnetic induction presents considerable difficulty to the teacher who wishes to give his students an accurate understanding of the fundamental principles of electromagnetism. This is partly because the subject involves inherent difficulties, but perhaps more because the usual simplified formulations are inadequate. As recently as 1922 the National Research Council published a bulletin[1] on electrodynamics, in which special attention was given to problems that were obscure because of the attempt to analyze them by simplified methods. Within the past year the matter has been brought up again in the pages of *Nature*.[2] The desirability of treating this subject correctly has already been discussed in this journal by Page and Adams,[3] and the objects of the present paper are to present a general formulation, to illustrate it by means of a few examples, and to point out some of the difficulties that may arise with the more customary treatments.

To all of the fundamental questions there can be given a precise answer in terms of the relativistic formulation of electrodynamics, but textbook writers have been slow to incorporate these results into their work. Presumably the matter has been regarded as too difficult, but it is very questionable whether difficulty is avoided by presenting students with inadequate concepts and methods if better ones are at all available. For a student who continues with the subject, it is probably as easy to learn the matter correctly in the first place, as to have to relearn it at some later stage; and for those students who find cultural values in the study of physics, an appreciation of the generality of the correct concepts is certainly of importance. A subject is correctly understood only when the concepts used are adequate for the treatment of all possible cases.

As pointed out by Page and Adams, the law of electromagnetic induction is often stated in two parts. One part comes directly from one of the Maxwell equations,

$$\mathbf{curl}\,\mathbf{E} = -(1/c)\partial\mathbf{B}/\partial t. \qquad (1)$$

It applies immediately to cases in which there is a rigid circuit in a varying magnetic field, and it states that the electromotive force is proportional to the rate of change of the total induction through the circuit. Problems involving this part of the law usually present little difficulty.

The other part of the law is not directly connected with any one of the Maxwell equations, but depends upon the proper interpretation of the whole set. It is often stated that a conductor "cutting lines of magnetic induction" has induced in it an electromotive force proportional to the rate at which the lines are being cut. This provides a visualization of the process; the student can picture the lines of induction in the space through which the conductor is moving, and can imagine that they have to be cut in order that the conductor may get through. Such lines of inductions have a definite location, and in order to fix this location they must be attached in some way to the physical apparatus. If, then, the apparatus, magnet or solenoid, is moved or rotated, the picture requires that the lines of induction move correspondingly. This picture leads to incorrect results in some cases, and it is only when the proper specification is made as to when the lines do and do not move that the formulation is correct. Such a specification destroys the pictorial value of the formulation and reveals it as merely a slightly clumsy means of stating the more general point of view.

Another formulation, the one emphasized by Page and Adams, is that a motional electromotive force is induced in a conductor that is moving in a magnetic field. It is based on the appreciation of the fact that an electric or a magnetic field is not a rigid body, and can move only in the sense that a wave moves. The field vector at any point can change with the time, and the distribution of the field can move, but the statement that the field

[1] *Electrodynamics of Moving Media*, Part II by John T. Tate, Bull. National Research Council, Vol. 4, Part 6, 1922.
[2] C. V. Drysdale, Nature **141**, 254, 907 (1938).
[3] Page and Adams, Am. Phys. Teacher **3**, 51 (1935).

moves may well be accompanied by incorrect connotations. In this formulation, the idea of "cutting" is replaced by the more precise idea of "moving in" or "moving through." *But the same difficulty remains;* it is necessary to specify the reference frame with respect to which the conductor is moving. This can be done, of course, but the process of doing it leads directly to the more general formulation.

A general formulation of the law, including both parts and applicable to any case, can be set up by recognizing the fact that electric and magnetic fields depend not only on the point in space and the instant in time to which they refer but also on the system of coordinate axes with reference to which they are expressed. It may be objected that much of the physical reality of the fields disappears if they are not independent of the system of axes, but it is just this kind of reality that such fields do not have. If the Maxwell equations are valid for one system of axes, they are valid for all systems moving uniformly with respect to the first; *but* the field quantities **E** and **B** at a given point and time are *not* the same when expressed with reference to the various sets of axes.

The transformation of fields from one system of axes to another is given in the special theory of relativity. However, when the velocity is very small, an approximate transformation can be derived by elementary means. If the electric field is defined as the force on a unit charge that is at rest with respect to the system of axes involved, it follows at once from the force equation that

$$\mathbf{E}'=\mathbf{E}+(\mathbf{v}/c)\times\mathbf{B}. \tag{2}$$

Here **E** and **B** are the electric field and the magnetic induction with respect to one system of coordinates, and **E**′ is the electric field in the system of axes that is moving with the velocity **v** with respect to the first. If, then, it is required that the Maxwell equations be invariant, in the same approximation, to the Galilean transformation between these systems of axes, it follows that

$$\mathbf{B}'=\mathbf{B}-(\mathbf{v}/c)\times\mathbf{E}. \tag{3}$$

By means of these transformations, if the electromagnetic field is known with reference to one set of axes, it can be found with reference to any other set moving uniformly with respect to the first. Eqs. (2) and (3) constitute, of course, only an approximate transformation, valid as $\mathbf{v}/c\rightarrow 0$; but the concept of a field that depends upon the system of coordinate axes with reference to which it is expressed is a precise concept, and one that is fundamental in much of electromagnetic theory.

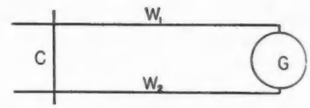


FIG. 1. Complete circuit in an infinite uniform magnetic field.

A general law of electromagnetic induction can then be formulated in the statement that the electromotive force around any circuit is given by $\int\mathbf{E}\cdot\mathbf{dl}$, where **E** is expressed, at every point, with reference to the system of axes moving with the conductor. This includes all aspects of the law of induction. When the circuit involved is rigid and the field is changing, it follows from Eq. (1) that the integral of the electric field around the circuit is proportional to the rate of change of the total flux. For the case of a simply connected circuit of linear conductors that is being deformed in a constant magnetic field, Eqs. (2) and (3) lead directly to the same result. In cases involving sliding contacts, the law requires that different parts of the integral be taken with respect to different systems of axes. Open circuit problems can be treated by using the requirement that the field inside a conductor, expressed with reference to the system of axes in which the conductor is at rest, be zero. The analysis of a number of examples will show the general applicability of this formulation as well as the special treatments needed when the other formulations are used.

As a first example, consider the arrangement indicated in Fig. 1. Here W_1 and W_2 are two parallel wires connected at one end by the galvanometer G; and C is a connecting wire that makes sliding contacts with W_1 and W_2, and can be moved along them perpendicular to its length. The whole system is in a uniform magnetic field **B**, normal to the plane of the figure, and extending indefinitely in all directions. This statement, that there is a uniform magnetic field and no electric field, presupposes the existence of a

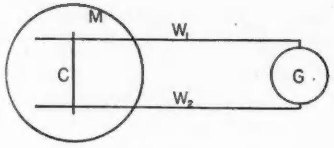


FIG. 2. Part of the circuit in a limited uniform magnetic field.

system of coordinate axes S_0 in which it is true. It implies that a system of axes S_0 can be set up with reference to which there is a uniform magnetic field and no electric field. Several experiments can be performed with this arrangement.

Exp. 1A.—Everything is at rest in S_0 except C, which is moved along the wires W_1 and W_2. The galvanometer shows a current, and the phenomenon can be described in all three ways:

(i) It may be said that the conductor C is cutting the lines of induction, and that on this account there is an e.m.f. induced in it. This seems an adequate statement in this case, although it must be implicitly understood that the lines of induction are fixed in the system S_0. Since the field is infinite, there is no apparatus to which they can be thought of as attached.

(ii) The conductor C is subject to a motional e.m.f. since it is moving in a magnetic field. It is rarely stated with reference to what system of axes this motion must exist, but it is clearly with reference to S_0.

(iii) There is an electric field parallel to C in the system of axes with respect to which C is at rest. There is no electric field with respect to the system S_0, and hence the e.m.f. is given by the integral along C. In this case it is also possible to speak of a change of the total flux through the circuit as causing the electromotive force.

Exp. 1B.—The connector C is held at rest in S_0 and the remainder of the circuit is moved. Again the galvanometer shows a current, and the phenomenon can be described in all three of the preceding ways. The electric field now acts on the wires connecting G with W_1 and W_2.

Exp. 1C.—The whole circuit is moved parallel to the wires W. No current is indicated by the galvanometer. Again all three of the modes of description are applicable:

(i) Conductor C and the wires to G are cutting lines of induction in such a way that the induced electromotive forces oppose and annul each other.

(ii) Both conductor C and the wires to G experience motional electromotive forces that annul each other.

(iii) Conductor C and the wires to G are in the same uniform electric field so that no current results. Here, also, the total flux through the circuit is constant.

The existence, in this arrangement, of a field stretching to infinity makes it impossible to give any meaning to the idea of moving the field. Such a situation could be defined only by the statement that there exists also an electric field. Cases of moving fields will be considered in the next example, which is based on the apparatus indicated in Fig. 2. Here everything is the same as in the previous case except that M represents the boundary of a uniform magnetic field. This field need not be considered as extending to infinity in the direction perpendicular to the section shown, for the magnetic circuit can be thought of as closed in some region outside the figure. However, the electric circuit and the magnetic field intersect only as shown. The system of axes S_0, with respect to which there is only a magnetic field, is the system with reference to which the magnet is at rest. Presumably the apparatus for supplying the current and, if desired, the walls of the laboratory are at rest in the same system of axes. If it is desired to make this quite definite, it may be specified that the whole apparatus is surrounded by a conducting shield, which establishes the system S_0.

Exp. 2A.—Everything is at rest in S_0 except C, which is moved along the wires W. The galvanometer shows a current, and the phenomenon is describable by any of the three methods:

(i) The conductor is cutting the lines of induction so that an e.m.f. is induced in it. With this arrangement there can be no question but that the lines of induction are fixed in S_0, since they must be attached to the apparatus that produces the magnetic field.

(ii) Conductor C experiences a motional e.m.f. because it is moving in a magnetic field, that is, with respect to the system S_0 and in a magnetic field.

(iii) There is an electric field in that system of axes with reference to which the conductor C is at rest, and none in S_0. The e.m.f. is then given by the integral of this field along C. The equivalent statement that the flux through the circuit is changing gives the correct results in this case.

Exp. 2B.—The conductor C and the magnet M are fixed in S_0. The rest of the circuit is moved. The galvanometer shows no current:

(i) The wires connected to G are not in a magnetic field and none of the wires are cutting any lines of induction.

(ii) The wires connected to G are moving but not in a magnetic field, so there is no motional e.m.f.

(iii) No electric field exists for the systems of coordinates with reference to which the various wires are at rest, at the

points where the wires are located, and parallel to their lengths. In this case there is no change of flux through the circuit.

Exp. 2C.—The whole circuit is moved, parallel to W, with reference to S_0. The galvanometer shows a current, and the phenomenon can be described in all three ways.

Exp. 2D.—The whole circuit is fixed and the magnet is moved along the direction of the wires W. This means that the shield surrounding the system is moved also. S_0 is now the system of axes in which the magnet is fixed and in which the field is purely magnetic, but the system with reference to which the electromotive force must be computed is the system S in which the wires are at rest. A current will exist in the galvanometer and the various descriptions are:

(i) The wire C is cutting the lines of force that move over it and experiences an e.m.f.

(ii) The wire C is not moving with respect to the observer, but it is moving with respect to S_0 and, hence, may be said to experience a motional e.m.f.

(iii) With reference to S_0 there is a pure magnetic field at C but none at G; hence, in the system S, in which the circuit is at rest, there is an electric field at C but none at G. This gives rise to an e.m.f. Also, the total flux through the circuit is changing.

Exp. 2E.—The magnet M and the conductor C are moved together and the rest of the circuit is held stationary. No current exists. Here again the observations are not made in the system of axes S_0 in which the field is purely magnetic, but in another system S moving with reference to it. In this system W_1, W_2 and G are at rest. This is merely Exp. $2B$ described for a different set of axes, but it is desirable to see that either set is suitable for its description:

(i) The wire C is not cutting lines of induction, for it must be understood that the lines are moving with the magnet, that is, are in system S_0.

(ii) The wire C is moving in a magnetic field, but it is not moving with reference to S_0 so it experiences no motional e.m.f.

(iii) In the system S_0 in which C is at rest there is no electric field. Neither is there any electric field at G in the system in which G is at rest. Also, the total flux through the circuit is constant.

The above experimental arrangements have involved closed circuits. Some problems of interest involve open circuits and a redistribution

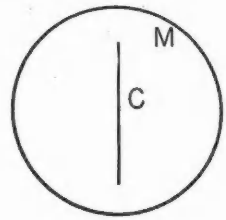


FIG. 3. Single conductor in a limited uniform magnetic field.

of charge due to the motion. Consider the arrangement in Fig. 3, where M represents the limits of a magnetic field, as in Fig. 2, and C is a wire which is to be moved perpendicular to its length. To determine whether charges accumulate at the end of the wire during an experiment, the wire may be thought of as broken during the motion and the charge of each part examined after the parts are brought to rest. The magnitude of the charge can be increased by connecting the ends of the wire to the plates of a condenser. Experiments of this type have been performed by various persons.[4]

Exp. 3A.—The wire is moved perpendicular to its length and a separation of charge is observed:

(i) The wire C is cutting the lines of induction and the charge separates until the e.m.f. is balanced by the static field.

(ii) The wire C is moving in a magnetic field so that it experiences a motional e.m.f.

(iii) It is now obviously impossible to speak of a change of flux through the circuit since there is no circuit; but since there is a pure magnetic field in the system S_0 in which the magnet is at rest, there is an electric field along the length of the wire in the system S in which it is at rest. The charge must then separate until this field is balanced.

Exp. 3B.—The wire is held fixed and the magnet is moved. A separation of charge occurs:

(i) If the lines of induction are regarded as moving with the magnet, the wire C is cutting them and experiencing an e.m.f.

(ii) The wire C is not moving with reference to the observer, but it is moving with reference to the magnet. This gives rise to a motional e.m.f.

(iii) The system S_0 in which the magnet is at rest is the one in which the field is purely magnetic. In the system in which the wire C and the observer are at rest, there is an electric field which produces a separation of charge in C.

The foregoing illustrations have involved only translational motion, whereas many of the usual

[4] See S. J. Barnett, Phys. Rev. **12**, 95 (1918).

experiments involve rotations. This introduces some additional complication since the Maxwell equations are not invariant to a transformation into a rotating set of axes. For this reason a definite meaning can be given to the idea of a rotation, and fields can be calculated in the usual way only for systems of axes that are not rotating. Although the special theory of relativity is not sufficient to handle these problems, they can be treated by means of the general relativistic formulation of the electromagnetic equations.[5] The significance of this formulation has been illustrated recently in an illuminating paper by L. I. Schiff.[6] Although the complete treatment of such cases involves the general theory of relativity, an approximate treatment for low velocities can be worked out as before.

The problem can be divided into two parts. The first is to find a system of axes for which the fields are known, and the second is to transform these fields into systems of axes moving with the various conductors. The transformation is not difficult. It is merely required that the motion at each point be approximated by a translation, and that the previously mentioned transformations—Eqs. (2) and (3)—be applied. Of course, these equations cannot be applied to points too far from the axis, for the translational velocity must be low. Neither can they be used right on the axis, since the direction of the motion must be definite. The first part of the problem is less simple. If it is possible to calculate the field in a system of axes that is not rotating, the Maxwell equations can be applied in the usual way, but for rotating systems other considerations must be introduced. In particular, it is sometimes convenient to surround the entire apparatus with a conducting shield. The method can be illustrated in a few cases.

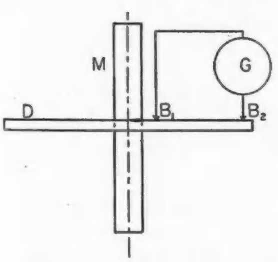


FIG. 4. The Faraday disk.

Consider the arrangement in Fig. 4, where M is a permanent magnet, or a solenoid carrying a current, and D is a conducting disk. Both the disk and the magnet can be rotated independently on their common axis. B_1 and B_2 are sliding contacts connected with the galvanometer G. In the discussion of the experiments let S_0 be the system of axes that is not rotating—that is, not rotating with respect to the primary inertial system of fixed stars—and let S be the rotating system of axes.

Exp. 4A.—The magnet M and the circuit are fixed in S_0, and the disk D is rotated. A current appears in the galvanometer:

(i) The conductor—the disk—is cutting the lines of force, insofar as such a motion can be referred to as "cutting."

(ii) The conducting disk is moving in the magnetic field and hence is subject to a motional e.m.f.

(iii) The total flux through the circuit appears constant, but, because the circuit is not composed of linear conductors, this is immaterial. Since in the system S_0 there is a pure magnetic field, at each point in S there will be a radial electric field proportional to the product of the magnetic field in S_0 and the velocity of the disk. This leads to an e.m.f.

Exp. 4B.—The magnet and the disk are rotated together, and the galvanometer shows a current:

(i) It is difficult to see how the disk is cutting the lines of induction since one might expect them to move with the magnet. However, it may be specified arbitrarily that the lines of induction are fixed in the system S_0 and do not partake of the motion of rotation of the magnet. On the other hand, if the lines are treated as rotating, they are cut by the rest of the circuit to give the e.m.f.

(ii) The disk is moving in the magnetic field, in the sense that it is moving with respect to S_0. It is then subject to a motional e.m.f. On the other hand, the rest of the circuit is moving with respect to S.

(iii) The change of total flux through the circuit is immaterial, so the fields in the systems of axes moving with the various conductors must be considered. It is a little difficult in this case to be sure in which system the field is purely magnetic. If the magnet is a coil of wire, its rotation will not change the current, and one might expect that the field will be purely magnetic in the system S_0. Then in the system S there will be a radial electric field, and this will lead to the necessary e.m.f. On the other hand, if it should be supposed that there is a pure magnetic field in the system S, the galvanometer circuit in S_0 will ex-

[5] See R. C. Tolman, ***Relativity, Thermodynamics, and Cosmology*** (Oxford Univ. Press). It may be questioned whether this formulation is sufficiently supported by experiment, but there is no experimental evidence against it.
[6] Schiff, Proc. Nat. Acad. 25, 391 (1939).

perience the electric field and provide the e.m.f. This experiment will not distinguish between these two possibilities, but because the Maxwell equations apply in S_0 but not in S, the first explanation is correct.

Various other combinations of motions can be devised, but they all lead to results understandable either in terms of a pure magnetic field in S_0 and the corresponding electric fields in S, or the converse. However, an experimental arrangement that might be expected to distinguish between the possible interpretations of the preceding experiment is shown in Fig. 5. This is an experiment actually performed by Barnett[7] and others. M is a magnet and C is a wire connecting the two conductors P_1 and P_2 of a cylindrical condenser. The wire C is arranged so that it can be broken during the rotation of either the magnet or the condenser in order to measure the charge on the conductors.

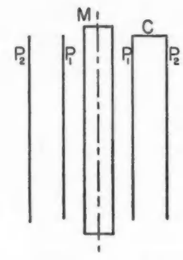


FIG. 5. Cylindrical condenser in the field of a bar magnet.

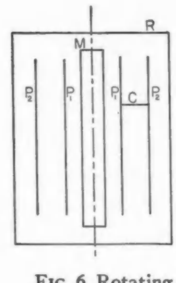


FIG. 6. Rotating system enclosed in a conducting shield.

Exp. 5A.—The magnet is fixed in S_0 and the condenser is rotated. The condenser becomes charged:

(i) The wire C is cutting the lines of force and hence an e.m.f. is induced in it.
(ii) The wire C is moving in a magnetic field and is subject to a motional e.m.f.
(iii) Since there is a pure magnetic field in S_0 there is a radial electric field in S and the condenser becomes charged until this field is balanced. There is no question of the change of flux through a circuit.

Exp. 5B.—The condenser is fixed in S_0 and the magnet is rotated. The condenser does not become charged:

(i) If the lines of force were to rotate with the magnet they would cut the conductor C and cause the condenser to become charged. Evidently the lines must be regarded as fixed in S_0.
(ii) The wire C is not moving and is not subject to a motional e.m.f.
(iii) In the system S_0 the field is purely magnetic so that no charging of the condenser is to be expected.

At first sight it would appear that Exps. 5*A* and 5*B* should give the same result, because the relative motions appear to be the same. The difference is in the motion of bodies at infinity. When the magnet is rotated, the walls of the laboratory must be rotated with it to give the equivalent situation. This can be approximated by surrounding the apparatus with a conducting shield R, as in Fig. 6.

Exp. 6A.—The magnet and the shield are fixed and the condenser is rotated. The condenser shows a charge. The descriptions are the same as for Exp. 5*A*.

Exp. 6B.—The condenser is fixed in S_0 and the magnet and shield are rotated together. Again the condenser charges up. Viewed from the system of axes S, this is the same as Exp 6*A*. Viewed from S_0 the descriptions are as follows:

(i) The shield is cutting the lines of induction which are fixed in S_0. This leads to a charge on the shield that induces a charge on the cylindrical condenser.
(ii) The shield is moving in a magnetic field and subject to a motional e.m.f. This produces a charge on the shield that induces a charge on the condenser.
(iii) In the system S the field is purely magnetic inside the shield because the charges on the shield produce an electric field in S_0. When the field in S is transformed to S_0 this electric field appears again and charges the condenser.

From the analysis of all these examples it appears that all three of the indicated methods can be used for the description of the phenomena if they are properly interpreted. The lines of force that are cut must be understood to be fixed in the proper system of axes, and the motional electromotive force is generated by motion relative to the proper system of axes. Nevertheless the determination of the proper interpretation involves in each case a reference to what is essentially the method of transformation of fields and destroys the pictorial value of the more usual methods. Hence it may be concluded that there is considerable merit in the use of the method of transformation of fields in the first place.

[7] Barnett, Phys. Rev. **35**, 323 (1912).

Reproductions of Prints, Drawings and Paintings of Interest in the History of Physics

8. A Display of the Arts and Sciences in 1698

E. C. Watson
California Institute of Technology, Pasadena, California

THE original of this reproduction, a magnificent engraving, $9\frac{5}{8}\times14\frac{7}{8}$ in. in size, was executed by Sébastien Le Clerc (see "Reproduction 7") in 1698 at just the time the Académie Royale des Sciences in Paris was being reorganized and was moving its collection of apparatus and its equipment to new quarters in the Louvre. It gives a remarkable conspectus of the state of physics at the beginning of the eighteenth century. It is considered to be Le Clerc's best work and is notable not only for the accuracy and fidelity with which the various scientific instruments are delineated but also for the number and variety of the subjects, their distribution and grouping, the handling of the lighting effects and the richness of the composition. As an engraving showing the state of scientific technic at the close of an epoch it has probably never been surpassed.

The Spring and Weight of the Air

Nora M. Mohler
Department of Physics, Smith College, Northampton, Massachusetts

Truth was thy aim, Experiments the way,
And Nature yielded to thy least Essay.[1]

An excursion into the writings of a scientist of the seventeenth century is an experience at once refreshing and illuminating, in its naïve yet scholarly approach to experimental science and in the light it casts on what constitutes the "scientific method" of our own times. In both respects the work of Robert Boyle is well worth the reading; his law concerning the relation between the pressure and volume of an enclosed mass of a gas at constant temperature can be proved by any first-year physics student in an afternoon, with apparatus that is simple but accurate. The procedure, computation and conclusion seem straightforward to the point of banality, and yet in all of his *New Experiments Physico-Mechanicall, Touching the Spring of the Air, and its Effects* published in 1660, there is no clear-cut statement of the law which we call by his name; and only under the prick of criticism and contradiction was this statement ultimately achieved. One finds very soon that ignorance of laws known to us, did not mean a corresponding mental vacuum or acknowledgment of lack of information, but that the new science of experimentation was kept busy trying to displace or modify general philosophic notions—of the "Vacuists" and "Plenists," of the followers of Aristotle and of Epicurus—as well as adding to the information gathered by Archimedes, Galileo, Pascal and the other early experimentalists. Perhaps the most important of these prevailing beliefs were that "nature abhors a vacuum" and that light objects rise because of their "positive levity," but there were many others. A theory of the atmosphere had to explain everything from hay fever to plagues, the vagaries of the weather and of the tides, the flash of meteors and the nature of combustion, and, in addition, had to be sufficiently plausible to displace other explanations, widely if not too critically accepted. This fog of misinformation seems to have been more of a hindrance than the prejudice against experimentation as such, and the lack of trained assistants and technical apparatus. But, in addition to its illumination on these points, this whole dissertation of Boyle's has a delightful personal flavor, severely removed from our current articles, so that there is revealed a gentleman of serious and religious bent, courteous and tender hearted, patient, persistent and indefatigable.

The volume was published in the form of a

Portrait of the Honourable Robert Boyle. Frontispiece of the *Works of the Honourable Robert Boyle*, edited by Thomas Birch. Engraved from a painting by T. Kersseboom. [Courtesy, Amherst College Library.]

[1] Elegy "On the Death of the Honourable Robert Boyle," a necrological sheet published in 1692. For a facsimile see J. F. Fulton, *A Bibliography of the Honourable Robert Boyle* (Oxford Univ. Press, 1932), p. 146.

letter to "my Lord of Dungarvan," and consists of the description of a long series of experiments, numbered and explained in meticulous detail. The reason for the publication he modestly ascribed to the persuasion of others:

> . . . intelligent persons in matters of this kind persuade me, that the publication of what I had observed touching the nature of the air, would not be useless to the world; and that in an age so taken with novelties as is ours, these new experiments would be grateful to the lovers of free and real learning.[3]

Then, after a bow of acknowledgment to his predecessor in such experiments—"that ingenious gentleman, Otto Gericke, consul of Magdeburg" —he proceeded to the business of the construction of apparatus: ". . . I put . . . R. Hook . . . to contrive some air-pump, that might not, like the other, need to be kept under water . . . and might be more easily managed." There follows a full and clear description of his, or rather Robert Hooke's, device: "that the person I addressed them to might, without mistake, and with as little trouble as possible, be able to repeat such unusual experiments."

Figure 1 shows the two essential parts of the apparatus, the glass vessel to be evacuated and the air pump itself. The vessel was no mean affair, but one of 30-qt capacity, fitted at the top with an air-tight plate and a "brass stopple *K*," which could be turned without letting in air, and holes through which strings for the manipulation of apparatus could be lowered; the bottom part of this vessel was drawn out into a neck fitted with a stopcock, and the glass, neck, stopcock and a tin plate were all cemented together. The "sucking-pump" below was mounted on a rigid 3-legged wooden frame, and had two main parts, a brass cylinder 14 in. long and 3 in. in diameter, and a sucker, also made of two parts. The first of these sucker parts was smaller than the cylinder, and was made air tight by a piece of tanned leather; fastened to this was an iron bar *5–5*, toothed on one side, smooth on the other. The cogs engaged in the cogs of a wheel, turned by a handle *7*. At the top of the cylinder was a valve *R*, consisting simply of a tapering hole with a fitted brass plug. In use, oil was added at the stopcock and on the leather of the piston, and the handle was turned until the sucker was at the top of the cylinder; then the valve was shut, and the sucker was pulled down; on opening the stopcock, air from the glass

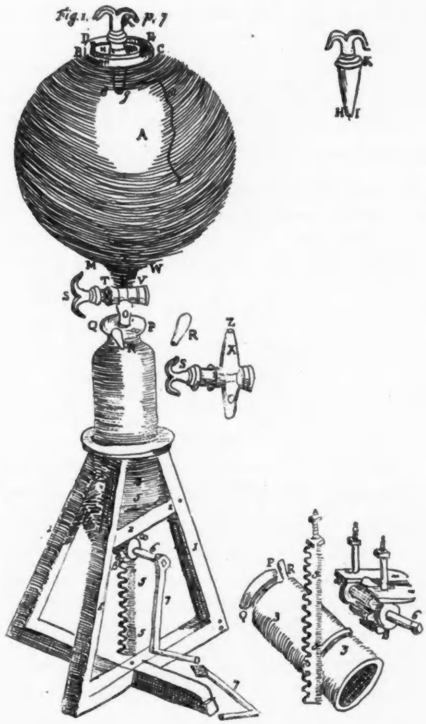


FIG. 1. Boyle's first air pump, showing the pump and glass vessel, as well as details of small parts. From the first plate of the Birch edition.

vessel expanded into the cylinder; after the cock was closed, the valve was removed and the operation repeated.

That Boyle was correct in his estimation of the interest his machine would arouse, is shown by the numerous changes made fairly soon in its design. The *Lexicon Technicon or, an Universal English Dictionary of Arts and Sciences*, by John Harris (1704), described this original form in detail, and others of quite different design, such as a double-action pump of Papin's, and an improved form of Boyle's in which we find the glass vessel replaced by a flat plate and a familiar looking bell jar. In the 1725 edition of Boyle's

[3] *The Works of the Honourable Robert Boyle*, ed. by Thomas Birch (1772), Vol. 1, p. 1. Unless otherwise noted, references are to Volume 1 of this edition.

works, the position of honor in the illustrations was given to a pump designed by Mr. Hauksbee, as an instrument "brought to its utmost degree of simplicity, and perfection," with the almost apologetic remark that "still there may be something left for future philosophers to do with it, besides repeating, varying and confirming his [Mr. Boyle's] trials."[3]

But no sooner had the construction of the apparatus been explained than the nature of the air was discussed in terms of the two distinct theories then current.

> For the more easy understanding of the experiments triable by our engine, I thought it not superfluous . . . to insinuate that notion . . . that there is a spring, or elastical power in the air we live in. . . .
>
> This notion may perhaps be somewhat further explained, by conceiving the air near the earth to be such a heap of little bodies, lying one upon another, as may be resembled to a fleece of wool. For this . . . consists of many slender and flexible hairs; each of which may indeed, like a little spring, be easily bent or rolled up; but will also, like a spring, be still endeavouring to stretch itself out again. . . .
>
> There is yet another way to explicate the spring of the air; namely, by supposing with that most ingenious gentleman, Monsieur Des Cartes, that the air is nothing but a congeries or heap of small . . . flexible particles, of several sizes, and of all kinds of figures, which are raised by heat (especially that of the sun) into that fluid and subtle ethereal body that surrounds the earth; and by the restless agitation of that celestial matter, wherein those particles swim, are so whirled round, that each corpuscle endeavours to beat off all others from coming within the little sphere requisite to its motion about its own centre . . . their elastical power is not made to depend upon their shape or structure, but upon the vehement agitation, and . . . brandishing motion, which they receive from the fluid aether.[4]

This compression and corresponding expansiveness is due not simply to the external force applied by means of the air pump but to the fact that "our atmosphere is a heavy body, and that the upper parts of it press upon the lower." "And though the height of this atmosphere, according to the famous Kepler, . . . scarce exceeds eight common miles; yet other eminent and later astronomers would promote the confines of the atmosphere to six or seven times that number of miles. . . . So that . . . a column of air, of many miles in height, leaning upon some springy corpuscles of air here below, may . . . bend their little springs, and keep them bent." There is in this connection some discussion of the idea of "positive levity," but in general Boyle's concern was more with a series of experiments to demonstrate and to measure this dilatation. One tiny bubble was watched through a series of "exsuctions," "so that the bubble of air taking up the room but of one grain in weight of water . . . by its own ἐλατήρ was so rarified, as to take up one hundred fifty-two times as much room as it did before." The breaking strength of this force was tried with glass alembics and round glass bubbles, with rather shattering results as far as the spheres were concerned. But then an unfamiliar note is sounded:

> And to evince, that these phaenomena were the effects of a limited and even moderate force, and not of such an abhorrency of a vacuum, as that to avoid it, many have been pleased to think, that nature must, upon occasion, exercise an almost boundless power; we afterwards purposely tried this experiment with several glasses somewhat thicker . . . and found . . . that it would not succeed; for the glasses were taken out as entire as they were put in.[5]

This is by no means the last that will be heard of the "horror vacui;" the subject keeps cropping up, each time as an untenable notion, but never considered settled. That Boyle knew his audience is shown by the fact that his experiments were explained in its terms by Franciscus Linus, and it was criticism and contradiction on this point that called forth the supplements so important in the history of the proving of "Boyle's Law;" of these, more later.

In the first nine experiments the general description of the way in which the apparatus and the air functioned mechanically was completed, and in experiment 10 the study of the effect of a vacuum on flame and on certain forms of life was initiated. Pages of inconclusive and sometimes contradictory results follow. The absence of a tenable theory concerning the nature of combustion made it impossible for Boyle to isolate the phenomena in which he was primarily interested, and so in the discussions we find our-

[3] *The Philosophical Works of the Honourable Robert Boyle, Esq.*, ed. and abr. by Peter Shaw (1725), Vol. 2, p. 406.

[4] Reference 2, p. 11.

[5] Reference 2, p. 25.

selves in a maze of mechanical-chemical-physiological puzzles on the nature of air, of fire and of life. His first experimental concern was with the problem of making a candle burn long enough in the apparatus after the air was removed "to discover, whether the extinction of the fire in the match did proceed from want of air, or barely from the pressure of its own fumes." So various other devices were tried, such as a burning-glass to fire some "dry and black" matter, "experience declaring things of that colour to be most easy kindled;" then live coals were studied and their variation in brightness with quantity of air noted, and mechanical methods of firing gunpowder in the glass vessel were tried. There were many difficulties in technic to be overcome, one of which at least still rouses the sympathy of any physics student: "but notwithstanding all our care and diligence the external air got in so fast, that after divers trials we were fain to leave off the experiment." As a result of such trials the suggestion was made that air has some "subtler parts" which penetrate even such tiny holes as those in good plaster.

Quite naturally this led to the questions as to what is a vacuum and what is air. To obtain the former, must ethereal, as well as material, substance be removed? Trials with a lodestone in and out of a vacuum showed that whatever substance was responsible for the transmission of this magnetic force was present in the vacuum undiminished. Light, as well as magnetism, is transferred and these beams

> . . . which rebounding from the seen object to our eyes, affect us with the sense of it: and that either these beams are corporeal emanations from some lucid body, or else at least the light they convey doth result from the brisk motion of some subtle matter.[6]

But his reasonable affirmation that little space would be occupied if this subtle matter were gathered together "with no parts between" does not help in the further question as to what constitutes air. For instance, when small bubbles appeared in the tube of water in the receiver, were they really "parcels of air" or a "spirituous part of the water"? Is air "ingenerable and incorruptible" or not? Heat experiments in which water was turned to vapor are discussed, but it is pointed out that the water was only "divided by heat . . . into very minute parts, which meeting together . . . do . . . return into such water as they constituted before." Yet, if the criterion for the existence of air is the possession of spring, then this rarefied water is air. Confusion is worse confounded by the production of "fictitious air," evolved by the action of oil of vitriol on nails. This possessed permanent spring and expanded when warmed, but its effect on flame and life were complicated to say the least:

> Lastly, to the foregoing arguments from experience we might easily subjoin the authority of Aristotle, and of . . . the schools, who are known to have taught, that air and water being symbolizing elements . . . are easily transmutable into one another.[7]

But if the atomists, from Leucippus and Epicurus to "divers modern Naturalists" are correct, there seemed to be no reason why the small parts of water or many other substances might not be "so agitated or connected as to deserve the name of air;" or their parts might be altered so as to act as little springs, as silver may be made elastic by hammering. So he left the question as to the nature of bubbles with a remark on "the dimness of our created intellects (which yet of late too many so far presume upon, as either to deny or censure the Almighty and Omniscient Creator himself)."

The subject of combustion led to that of breathing, and this also Boyle investigated somewhat later. The inquiry started as a mechanical one—can an insect fly in a vacuum? But the bee which had been persuaded to light on a flower fell without even attempting flight—whether from the thinness of the air, or from weakness, he tried to discover. Experiments with a lark and then a mouse were followed by "A Digression containing some Doubts touching Respiration,"[8] in which the whole question was discussed from the point of view of the authority of Galen and Hipparchus, Paracelsus and Aristotle, and that of experiment, with the conclusion that "there is some use of the air which we do not yet so well understand." And then Boyle proceeded to the discussion of submarines:

[6] Reference 2, p. 37.

[7] Reference 2, p. 54.

[8] Reference 2, p. 99.

> . . . a conceit of that deservedly famous Mechanician and Chymist, Cornelius Drebell, who, among other strange things that he performed, is affirmed, by more than a few credible persons, to have contrived . . . a vessel to go under water; of which, trial was made in the Thames, with admired success, the vessel carrying twelve rowers, besides passengers. . . . Drebell conceived, that it is not the whole body of the air, but a certain quintessence (as Chymists speak) or spirituous part of it, that makes it fit for respiration; . . . besides the mechanical contrivance of his vessel, he had a chymical liquor, which he accounted the chief secret of his submarine navigation.[9]

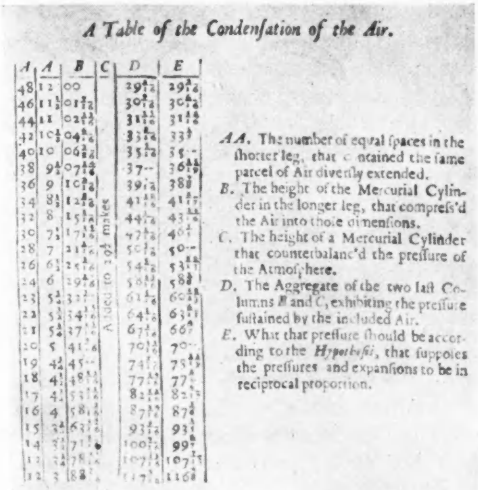
The discussion of the cherishing of the vital flame "continually burning in the heart" was followed by experimental work on puppies, on eels—which seemed to use the "little parcels of interspersed air" in water—on snails and bees and flies; then a final reluctant departure from this study with the conclusion that air may cool the blood, is more important in its "depuration," but with the suspicion "that the air doth something else in respiration, which hath not yet been sufficiently explained."

The production of "fictitious air" by dropping coral into spirits of vinegar was tried in a vacuum and found to be accompanied by "vast and numerous" bubbles. Then hot water was made to boil in the vessel by the reduction of the pressure, was slightly cooled and boiled again by further pumping out of the air, resulting in the production of "prodigiously vast bubbles." From this the conclusion was drawn that "pressure of the air might have an interest in more phaenomena than men had hitherto thought,"[10] particularly in the "operations of that vehement and tumultuous agitation of the small parts of bodies, wherein the nature of heat seems chiefly, if not solely, to consist."

But of most interest to the physicist is the work that eventually led to the statement of Boyle's law. The only experiments on this subject in the 1660 *Physico-Mechanicall* reports were repetitions of the famous Torricellian experiment, and they required no little skill; and to him, as well as to us, this is the "principal fruit" from his "engine."[11] A long slender tube of glass, sealed at one end, was filled with mercury, inverted into a vessel half-filled with mercury, and the whole lowered by strings into the receiver. The height of the mercury column was about 27 in. and this level was carefully watched as the pump was plied, and discussed on the assumption that the fall would be "in proportion to the exsuction of the air." It was noted how the fact that the level was maintained on shutting the vessel of mercury in the enclosure before pumping, showed that the mercury column was held up by the spring rather than by the weight of the air. The pumping was continued until scarcely an inch of mercury showed, and then air was let in bit by bit so that the mercury was lowered, impelled upward or checked at will. It was a famous experiment, repeated before various observers, including Doctor Wallis the mathematician, and Mr. Wren, better known to posterity as an architect than as a physicist. No exact hypothesis of the proportion of fall could be worked out, although the importance of such a generalization was realized; but Boyle acknowledged himself too little of a mathematician to undertake the calculation of the volumes of the receiver and the air displaced by the piston.

A similar barometer tube was set up in a room and its height observed for several months. There was some variation with temperature, attributed to the remnant of air in the top of the tube, but often an additional and unexpected rise or fall that could not be so explained. These changes were both puzzling and a nuisance; they were ascribed to "steams" rising from the earth, but they masked the effect sought, which was nothing more nor less than a crucial test of the two theories concerning the ebb and flow of the sea. According to Descartes, a greater pressure put by the moon upon the air and so transmitted to the sea was the cause of the variation in the height of the sea; but the assumption that the moon's attraction was the cause led to the opposite conclusion. The barometric variation masked such effects entirely as well as providing a new puzzle of "strange ebbings and flowings, as it were, in the atmosphere, or . . . great and sudden mutations, either as to its altitude or its density, from causes, as well unknown to us, as the effects are unheeded by us;" but this "new mystery of

[9] Reference 2, p. 107. For a further account see G. Tierie, *Cornelis Drebbel* (Amsterdam, 1932).
[10] Reference 2, p. 115.
[11] Reference 2, p. 33.



A Table of the Condenſation of the Air.

AA. The number of equal ſpaces in the ſhotter leg, that contained the ſame parcel of Air diverſly extended.
B. The height of the Mercurial Cylinder in the longer leg, that compreſs'd the Air into thoſe dimenſions.
C. The height of a Mercurial Cylinder that counterbalanc'd the preſſure of the Atmoſphere.
D. The Aggregate of the two laſt Columns B and C, exhibiting the preſſure ſuſtained by the included Air.
E. What that preſſure ſhould be according to the *Hypotheſis*, that ſuppoſes the preſſures and expanſions to be in reciprocal proportion.

FIG. 2. Data for the proof of "Boyle's law," with an explicit statement of that law given in "E." From "A Defence of the Doctrine touching the Spring and Weight of the Air" (1662), p. 60.

nature" was brought closer to the weather and to human understanding by comparison with the "pains and aches that are often complained of by those that have had great wounds or bruises, and that do presage great mutations in the air."[12]

Shorter series of experiments included many fields of physics; there were experiments on the lifting power of magnets in rarefied air and on the elusive flashes of light occasionally seen in the vessel and thought to be due to some "change in texture of the air." Trial of the period and length of time a pendulum would continue to vibrate "seem'd to teach us little," as the variations with and without air were slight. The work on sound shows both Boyle's ingenuity and his hesitancy in drawing definite conclusions.

> That the air is the medium, whereby sounds are conveyed to the ear, hath been for many ages, and is yet the common doctrine of the schools. But this received opinion hath been of late opposed by some philosophers upon the account of an experiment made by the industrious Kircher, and other learned men.[13]

In Kircher's experiment a bell was hung in the top of a barometer tube, and the clapper was swung by a magnet; as the sound could still be heard, a "more subtle body" than air was thought to be the medium of transfer. Boyle suspended a watch with opened case from the top of the receiver by a packthread "as the unlikeliest thing to convey a sound to the top of the receiver" and then plied the pump. The sound grew progressively fainter until "neither we, nor some strangers . . . could, by applying our ears to the very sides, hear any noise from within," and then louder as air was let in: "Which seems to prove, that whether or no air be the only, it is at least the principal medium of sounds." Instead of leaving this as settled, he continued, rather unfortunately, by suspending a small bell by means of a bent stick which pressed against both sides of the receiver; on the evacuation of the receiver they "could not discern any considerable change in the loudness of the sound" and modified the former conclusion by remarking that after all a more subtle matter may be the medium, or else that very few particles of air are sufficient. Other experiments were planned but postponed because suitable apparatus was not available. One of them concerns an attempt to make this "subtle medium" disclose itself; an ingenious bellows was to be depressed in the evacuated jar in such a way that a breath of ether would be blown into a tiny musical pipe or against a feather. The trial of this, like other more famous ether experiments, led to a null result. It is this later series of experiments, reported in *A Continuation of new Experiments Physico-mechanical, touching the Spring and Weight of the Air, and their Effects* (1682), that is now considered as definitely settling this question of air as the medium for the transfer of sounds.[14]

A few experiments on siphons are reported, with "little or no cause to doubt, . . . that the course of water through siphons depends upon the pressure of the air. . . . This occasion . . . puts me in mind of an odd kind of siphon," discovered by "some inquisitive Frenchmen" who observed "an unusual rise of water in slender pipes of glass." This Boyle had previously observed himself, but "presuming it might be

[12] Reference 2, p. 42.
[13] Reference 2, p. 62.

[14] D. C. Miller, *Anecdotal History of the Science of Sound* (Macmillan, 1935), p. 21.

casual, I had made little reflection upon it." There follows a brief but puzzled discussion of capillarity. He obtained a rise of 5 in., made a tiny siphon, noted that the height was increased by wetting the tube and unchanged by tilting the tube, and found little or no change in the evacuated vessel. The cause of this ascent seemed very difficult to explain, but he remarks that it would be pertinent to inquire why the surface of water shows a concave surface in a tube, whereas mercury shows a convex surface and a corresponding depression.[15]

More familiar in argument and technic was Boyle's work on the density of air. His methods and results are briefly and definitely set forth,[16] and where these were at variance with the classical experiments, so much the worse for the latter. An aeolipile was heated, stopped with wax, cooled and weighed; the wax was pierced and the weighing repeated; it was filled with water and reweighed with scales "that would turn with the fourth part of a grain." He compared his result—that "water is near 1000 times heavier than air"—with the results of Galileo, Mersennus and Ricciolus, to their disadvantage, and continued with a comparison of the weights of mercury and air, and finally of water, with the result in modern notation, of 13.8. This result he defended, "tho' the illustrious Verulam, merely for want of exact instruments, makes the proportion between them greater than 1 to 17." He calculated the height of an atmosphere of uniform density as 7 mi. Other experiments on the weight of air in air and of the action of siphons in a vacuum are included before the end of this long letter is reached.

This volume, *Physico-Mechanicall Experiments*, was published in 1660. As a whole it has been presented as evidence both for and against the influence of Francis Bacon, "my Lord Verulam," on the scientists of the period. It seems impossible for physicists to discuss the question without rancor. The Baconian method of listing all objects that display a certain quality, then all

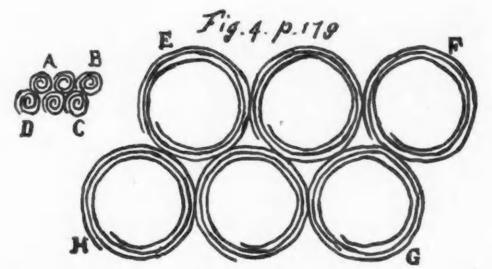


FIG. 3. The springs which cause the spring of the air! From the first plate of the Birch edition.

the objects that, although otherwise similar, do not display it, and from these two lists deducing the essential nature of that quality, seems so artificial in the light of quantitative experiments as to be worse than useless. Bacon himself applied the method to the problem of heat with the conclusion that heat is a "mode of motion;" but most readers who follow his argument are rather forced to the conclusion that the result was due more to luck than clear analysis. Any experiment, whether inconclusive or not, whether performed with imaginative insight or not, was considered a worthy addition to that sum total of data, which, when completed, would automatically lead to correct conclusions. It is interesting to note that whether Boyle was wholeheartedly a Baconian or not, he used this method in a discussion of the likenesses and differences between phosphorescence and burning coal,[17] and although no conclusions were drawn from the lists, neither were any arrived at on the basis of his descriptive experiments on the same subject. His respectful references to the "illustrious Verulam" have been explained away on the assumption that it was lip service to a prominent exponent of liberal endowments for laboratory equipment and workers; he speaks with equal respect of his experimentalist forbears, of Galileo and Pascal and von Guericke. It remains difficult to refrain from the conclusion that the example of such experimentation as Pascal's was more important in Boyle's scientific education than the *Novum Organon*.

The controversies which the publication of Boyle's book aroused among his contemporaries

[15] Reference 2, p. 80. Rohault in the *Traité de physique* (1671), Pt. 1, Chap. xxii, Sections 81–85, describes the phenomenon and explains it in terms of the impossibility of air pressure being effective in tiny tubes. Hauksbee gives the correct explanation in his *Physico-Mechanical Experiments* (1709).

[16] Reference 2, p. 86.

[17] "Observations and Tryals about the Resemblances and Differences between a Burning Coal and Shining Wood," Phil. Trans. No. 32, p. 605 (1668).

were of a different kind but not lacking in virulence. The attacks of Franciscus Linus and Thomas Hobbes concerning the notions of a vacuum and its possible existence are important to us chiefly in the refutation made by Boyle in a second edition of the *Physico-Mechanical Experiments*, published in 1662. Quotations from the attacks were given verbatim, followed in each case by a long discussion and reiteration of his original arguments and evidence. The objections raised by Linus were not that air has neither spring nor weight, but that these are too small to produce the effects ascribed to them. The foundation for his conviction seems to have been physiological. If the Torricellian experiment is repeated with a tube closed by a finger instead of sealed off, the flesh protrudes slightly into the evacuated space left at the top of the tube after the inversion into a cup of mercury, and there is also a feeling of tension in the flesh. This Linus ascribed to the existence of a "funiculus," a set of invisible and tenuous strings supporting the mercury and in some way fastened to the finger. Of course, the chief advantage of this theory was that it filled the vacuum with the funiculus, and so saved a centuries-old theory. Boyle's refutation seems largely wordy and repetitious, but the crux of it is refreshingly clear-cut. There is first a repetition of the famous experiment devised by Pascal, in which the difference in the height of a mercury column was measured at the foot, and on top, of a mountain with a device sensitive enough to register the change in level from the leads to the ground of the "lofty abby-church of Westminster." Then he came at last to "Two new Experiments touching the measure of the force of the spring of air compressed and dilated."

> We took then a long glass-tube, which, by a dexterous hand and the help of a lamp, was in such a manner crooked at the bottom, that the part turned up was almost parallel to the rest of the tube, and the orifice of this shorter leg of the siphon . . . being hermetically sealed, the length of it was divided into inches . . : by a streight list of paper, which . . . was carefully pasted all along it. Then putting in as much quicksilver as served to fill the arch . . . that the mercury standing in a level might reach in the one leg to the bottom of the divided paper, and just to the same height in the other. . . . This done, we began to pour quicksilver into the longer leg of the siphon, which by its weight pressing up that in the shorter leg, did by degrees streighten the included air: and continuing . . . till the air in the shorter leg was by condensation reduced to take up but half the space it possessed (I say, possessed, not filled) before; we cast our eyes upon the longer leg of the glass, . . . and we observed, not without delight and satisfaction, that the quicksilver in that longer part of the tube was 29 inches higher than the other. Now that this observation does both very well agree with and confirm our hypothesis, will be easily discerned by him, that takes notice what we teach; and Monsieur Paschal and our English friends' experiments prove, that the greater the weight is that leans upon the air, the more forcible is its endeavour of dilatation, and consequently its power of resistance. . . . For this being considered, it will appear to agree rarely-well with the hypothesis, that as according to it the air in that degree of density and correspondent measure of resistance, to which the weight of the incumbent atmosphere had brought it, was able to counterbalance and resist the pressure of a mercurial cylinder of about 29 inches, as we are taught by the Torricellian experiment; so here the same air being brought to a degree of density about twice as great as that it had before, obtains a spring twice as strong as formerly.[18]

The tube was broken, replaced, and the readings taken in the most professional style, with columns of figures of the height of the column of mercury, the corresponding volume of air, the atmospheric pressure, total pressure, and "what the pressure should be according to the hypothesis, that supposes the pressures and expansions to be in reciprocal proportion." Numerical discrepancies were present of course, but "the variations are not so considerable, but that they may probably enough be ascribed to some such want of exactness as in such nice experiments is scarce avoidable." There follows work on the rarefaction of the air, with a similar tabular representation of data. Still further work was planned "but being then hindered by some unwelcome avocations to prosecute those experiments, we shall elsewhere, out of other notes and trials (God permitting) set down some other accurate tables concerning the matter."[19]

In the other pages of refutation first of Linus' and then of Hobbes' attacks, perhaps the brightest spot is the description and diagram of "these small coyled particles of the air" of 10^{-12} in. diameter,[20] and their expansion by a factor of 10.

[18] Reference 2, p. 156.
[19] Reference 2, p. 161.
[20] Reference 2, p. 179.

Heat is pictured as causing a sort of centrifugal action, resulting in the uncoiling of each spring. But there is nothing even slightly comparable in importance to the experiments just described.

For a contrast to the care in experimentation, patience in explanation and wordiness of exposition on the part of Robert Boyle, it is interesting to turn to the *Treatise on the Weight of the Mass of the Air*, by Blaise Pascal, published in 1663, a year after his death, by his brother-in-law, M. Perier, who had so ably assisted him in his famous Puy de Dôme experiment. The distinctive quality of his style appears in the first sentence:

> It is no longer open to discussion that the air has weight. It is common knowledge that a balloon is heavier when inflated than when empty, which is proof enough. . . .
>
> This principle being laid down, I will now proceed to draw from it certain consequences.

and proceeds through the description of various experiments to:

> Let it then be set down, (1) that the mass of air has weight; (2) that its weight is limited; (3) that it is heavier at some times than at others; (4) that its weight is greater in some places than in others, as in [highlands and] lowlands; (5) that by its weight it presses all the bodies it surrounds, the more strongly when its weight is greater.[21]

The phenomena popularly supposed to be caused by nature's abhorrence of a vacuum are described and then their explanation in terms of fluid pressure is given. The argument is mainly by analogy with heavier liquids, such as tubes of mercury in water, and the application to these cases of the laws of liquid pressures is proved in the first treatise of this slim volume. There is no catering to your doubts. "This will be perfectly convincing." "These facts are so intelligible, easy, and simple, that it is strange that recourse should have been had to the abhorrence of a vacuum, occult qualities and other such far-fetched and chimerical causes for the purpose of accounting for them." But in spite of the clean clarity of his almost mathematical argument, doubts certainly creep in. The pistons that slide easily but are air tight, the long tubes of mercury lowered into a river and read there, start one wondering; and the fly that lives as easily in lukewarm water as in air, the man that perches placidly on a rock on the bottom of a river observing the "cupping effect" of the 20-ft tube he holds against his thigh—these do not bear comparison with the meticulously honest observations of Boyle. Evidently the line of argument was so clear to Pascal himself, especially as many of the experiments had been successfully tried, that to imagine one was the equivalent of its performance. The most famous of all was actually performed by M. Perier at Pascal's instance but without detailed directions, and it is to this experiment of 1648 that Boyle refers. Pascal requested his brother-in-law to compare the pressure of the atmosphere as measured by a mercury barometer at the foot of the Puy de Dôme with that at its summit, some 500 fathoms higher. The checks and counterchecks which made the trial so convincing seem due to Perier himself; one observer was left with a tube at the foot of the mountain to measure the height at intervals during the day. The second tube was filled and compared with it at the beginning and end of the trip, and readings were taken at intervals during the climb as well as at the top of the mountain. The total change due to elevation was found to be 3 in. $1\frac{1}{2}$ lines: "We were so carried away with wonder and delight, and our surprise was so great that we wished, for our own satisfaction, to repeat the experiment."[22] It is primarily from this work that Pascal concludes: "All these effects, whether of the weight of air or . . . of water, follow so necessarily from the equilibrium of fluids that there is nothing more evident in the world." Then "for the pleasure of it" he computed the weight of the whole mass of the air and found it to be[23] 8,283,889,440,000,000,000 lb. And, although Boyle may have had doubts as to Pascal's qualifications as an exponent of the experimental method when his fly drowned, Pascal himself had none: "let them learn that experiment is the true master that one must follow in Physics."

Although Pascal performed his task with so

[21] *The Physical Treatises of Pascal The Equilibrium of Liquids and The Weight of the Mass of the Air*, tr. by I. H. B. and A. G. H. Spiers, ed. by F. Barry (Columbia Univ. Press, 1937), p. 27.

[22] Reference 21, p. 105. From a copy of a letter sent by Monsieur Perier to Monsieur Pascal the Younger, September 22, 1648.

[23] Reference 21, p. 66.

much more economy than Boyle, it is to the latter that we turn for a fascinating picture of the ferment of scientific and pseudo-scientific ideas prevalent in the seventeenth century. The refutation of the philosophic ideas held since the time of Aristotle, of the abhorrence of a vacuum and the quality of levity, is mingled with the discussion of rival theories of light and heat and the structure of matter. The importance in our own thinking of the existence of simple theories in other sciences, the technic concerning the isolation of pertinent facts, the classification and mathematical analysis of data, as well as a body of experimental technics, are seen to be assumed concomitants of our scientific method. Perhaps it is well once in a while for the physicist of the twentieth century, involved in such complexity of theory and plethora of phenomena that it is difficult to see the woods for the trees, to share for a few minutes the confusion of his forbears who could not see the trees for the forest until a vast amount of clearing and trimming had been done. And certainly to Robert Boyle, the serious and charming virtuoso, in whom are found the fumblings incident to the transition from descriptive to quantitative reports and a steady effort toward the clarification of the statement of theories, we owe a very considerable debt which it is a pleasure to acknowledge.

Physics Survey Courses *Versus* Physical Science Survey Courses as Agencies of General Education

WILLARD GEER
Long Beach Junior College, Long Beach, California

CURRENT professional literature has been calling attention more or less frequently to the subject of "survey" courses, their status, their objectives, their merits and their place in the curriculum. In an effort to pool the experiences of the instructors in such courses in physical science, the author has undertaken a comprehensive study of physical science and physics survey courses as they exist at the present time in the colleges and universities, teachers colleges, and junior colleges of the United States. In order to limit the study to workable dimensions, institutions for special races, sexes or professions (except teaching) were omitted from the study. The entire investigation included 868 institutions, but the principal data were obtained from an intensive study of 85 physical science survey courses and 38 physics survey courses which represented a wide range in institutional size and type, and geographic location.

The present article describes that portion of the complete study which deals with a comparison of the practices in use in the two types of courses. Before proceeding to the actual contrast between the two types, the definitions used for their selection should be stated.

A *science survey course* is defined as a course offered for general education or for orientation that cuts across the fields of the separate sciences and is, for the most part, cooperatively administered by several departments or by a group of teachers from several science fields. A course in which semesters or parts thereof are devoted to the separate sciences, but in which the work is so integrated as to present a real survey picture of two or more of the sciences, is also considered a science survey course. A *physics survey course* is considered to be any course in physics of elementary collegiate level that has as its main purpose a contribution to general education—a course for the layman in contradistinction to those technical and highly mathematical approaches that are designed for future specialists in science; a course that presents a broad overview of the entire field of physics rather than precise experience with the details of the subject.

Throughout the article, where survey courses are to be compared with other physics or science

courses offered at the same level, the terms *academic*, *regular*, *classical*, *traditional* and *pre-professional* have been used interchangeably to describe the courses that are ordinarily offered for the student who intends to specialize in science or to take further sequences in it.

Perhaps the most graphic way to compare these two types of survey courses is to assume that the experience of the majority of teachers cooperating in this study defines a typical course of the one type or the other as it exists today, and upon this assumption to "paint" composite pictures of these courses side by side. It must be borne in mind, however, that the practices and procedures here reported are not presented as either "good" or "bad," since they have never been subjected to any valid experimental testing for evaluation. They are presented as representative empirical standards, with a hope that they will open up the way for some much needed experimental work on their values.

I. Status

Typical Science Survey Course	Typical Physics Survey Course
1. This course is offered in 1 out of 6 universities and colleges, in 1 out of 2 teachers colleges and in 1 out of 10 junior colleges.	1′. This course is offered in fewer institutions; it exists in 1 out of 9 universities and colleges, in 1 out of 17 teachers colleges and in 1 out of 36 junior colleges.
2. In teachers colleges half of these courses have been established within the past $3\frac{1}{2}$ yrs; in the other types of institutions they have been established, for the most part, within the past 2 or 3 yrs.	2′. Half of these courses have come into the curriculum within the past 4 yrs.
3. The instructor in charge believes that this course has found a permanent place in the curriculum of his institution, but that certain aspects of its organization are still in a state of experimentation.	3′. The instructor in charge believes that this course has found a permanent place in the curriculum of his institution.
4. In universities and colleges, the mean enrolment is 100 students per year, with 16 percent of the institutions reporting more than 200 in the class; in teachers colleges, the mean is 130, with 25 percent reporting more than 200; and in junior colleges the mean is 144 students per class, with 17 percent reporting over 200. The ranges for the three types of institutions are 10–700, 10–675 and 8–825, respectively.	4′. The enrolment in universities and colleges averages 82, with 8 percent enroling more than 200; teachers colleges average 35; and junior colleges, 14. The ranges are 3–500, 24–45 and 12–18, respectively.
5. The enrolment has increased approximately 10 percent in the past 2 yrs. The increase is larger than this in universities and colleges and smaller in teachers colleges where the enrolments were already large.	5′. The enrolment has increased 15 percent during the past 2 yrs.
6. This course is one out of the many in which the present total enrolment may be conservatively estimated as 25,000 students.	6′. The present total enrolment may be estimated conservatively at 4500 students.
7. The chance is even that this course, if in a university, college or junior college, is offered on an elective basis. If it is in a teachers college, the chance is 6 to 1 that it is required.	7′. This course is very probably offered on an elective basis, regardless of the type of institution in which it is found.
8. It is a lower division course and, if found in an institution where it is required, is probably given in the freshman year.	8′. It is a lower division course.
9. The catalog title is probably "Science Survey," "Physical Science" or some similar title.	9′. The catalog title is probably "Physics Survey" or some very similar title.
10. It is a course of 2 semesters duration, each semester bearing 3 hours credit. The range in reporting schools is 2 to 12 hours credit, with 25, 27 and 50 percent of universities and colleges, teachers colleges and junior colleges, respectively, offering the course in 6 semester hours. About 60 percent of the universities, colleges and teachers colleges, and 80 percent of the junior colleges ask 2 semesters for this course.	10′. Same.
11. Students may enter without prerequisites except such as admit them to the institution.	11′. Same.

12. Students may not enter *second-year* pre-professional courses in science after having had only this course as preparation.	12′. Same.
13. If a student intends to major or minor in physics, he may possibly be barred from the course.	13′. If a student intends to major or minor in physics, he most probably will be barred from the course.
14. The chances are even that the course has been designed especially for freshmen, but the enrolment is not confined to them.	14′. It is not confined to freshmen, and probably has not been designed especially for them.
15. However, the class is made up principally of freshmen.	15′. Less than half the students are freshmen.

II. Impetus for Establishment of These Courses

1. The course was initiated either by the administration or by the science department.	1′. It was started probably by the physics department, but in some instances may have been initiated by the administration.
2. It was established because of (1) the failure of the regular pre-professional or academic courses in science to meet the needs of the majority of the students in the matter of broad cultural background, and (2) certain requirements imposed by the state department of education, some teacher-certification requirement or some local institutional requirement.	2′. The reasons for establishment were: (1) same, and (2) because of special local requirements calling for such a course.
3. This course, rather than the *physics survey* course, was chosen because of the feeling that students do not have sufficient time in college to devote to separate courses in all the science fields, and because separate courses would defeat the purpose for which the survey course was added to the curriculum.	3′. This course was chosen in preference to the *science survey* course because of the feeling that the material which should be taught could not be covered adequately if the other sciences were included.

III. Aims and Objectives

1. Among the objectives considered most important are "understanding of fundamental scientific laws and principles," "acquisition of general scientific knowledge" and "ability to explain natural phenomena." In universities, colleges and junior colleges, one of the most important aims is "freeing from superstitions, groundless fears and quackery."	1′. The objectives seem to be identical with those governing the *science survey* courses.
2. While objectives concerned with use of the "scientific method" are secondary to those just mentioned, they are considered important.	2′. Instructors here are less concerned with the objectives relating to use of the "scientific method."
3. Objectives common in pre-professional courses in science or those that tend to develop specialists, or that are concerned with developing vocational or avocational interests in science are not considered important here.	3′. Same.
4. Objectives relating to the social implications of the work of scientists or to the need for continued research are not emphasized. Nor is such an objective as "knowledge of the history of science" considered important.	4′. Same.

IV. Subject Matter and Its Treatment

1. In the choice of appropriate subject matter, the instructor in charge used for a criterion his own judgment or the material contained in the adopted text.	1′. The instructor most probably used only his own judgment in the selection of subject matter, although in some cases he was governed somewhat by the material in the adopted text.
2. The course tends to be[1] "selective," rather than "encyclopedic," in choice of material. It tends to be slightly more "interpretative" in presentation than to present "science for its own sake." Relationships between the subject matter and that from other fields of knowledge are stressed. The instructor seeks the middle ground between organizing his subject matter	2′. This course tends to be less strongly "selective"; it tends a bit more to the middle path between the highly "interpretative" and the presentation of "science for its

[1] See Havighurst, Am. Phys. Teacher 3, 97 (1935).

in the "logical order" of classical physics and organizing it around the "problem of human needs." He attempts a presentation that is divided between pure theory and practical application of theory. He seeks a presentation that is "stirring" and intended to bring out an emotional reaction in behalf of science and its work.

own sake." Like the other course, it stresses relationships. The instructor tends to arrange his material more in the "logical order" of classical physics than does the teacher of the other type of course, but, like him, seeks a middle path between pure theory and practical application of theory. He differs strongly from the *science survey* teacher in the matter of arousing emotions, and emphatically insists upon an entirely objective presentation.

3. The order of decreasing time-emphasis given to various fields is: physics, chemistry, astronomy, geology and meteorology.

3′.

4. In the physics portion of the course, electricity and magnetism are given the most time, light next, then mechanics, heat, properties of matter, sound, modern physics, and history and appreciation, in the order named.

4′. In time-emphasis, this teacher agrees in placing electricity and magnetism first, and history and appreciation last. But for the rest he differs, using the following order: mechanics, light, heat, modern physics, sound and properties of matter.

5. Throughout the physics portion of the course, the same topics are stressed as in academic physics, but the topics that are new to the field of physics or that are practical daily applications of physical principles are given proportionately less consideration. This tendency is not quite so strong here, however, as it is in the *Physics Survey* course.

5′. The teacher stresses almost the same topics in the separate divisions of physics as does the *science survey* teacher, but tends still more strongly to parallel the stress of academic physics.

6. Demonstrations are given rather widely throughout the entire range of topics in each division of physics, their frequency of occurrence in the various divisions being in descending order: light, electricity, mechanics, modern physics, sound and heat. But again, topics of academic physics are stressed by demonstration at the expense of the newer or more practical aspects of the subject, except where ease or difficulty of demonstration is a significant factor.

6′. No significant difference.

7. In the matter of mathematical stress of the various topics, the course is characterized by more problems in mechanics and heat than in any of the other divisions of the subject. Very few problems in modern physics are required. In electricity and magnetism, the "concepts of fundamental units of electrical measurement" and "effect of currents" are illustrated with mathematics. In light, "lens problems," "reflection" and "refraction" appear with some mathematics. In mechanics, "work, power, energy and machines," "accelerated motion," "gravitational field," "statics" and "hydraulics" are given mathematical illustration. In heat, "temperature scales," "calorimetry," "expansion," "general gas laws" and "transmission of heat" have mathematical approaches. In sound, "organ pipes" and "wave equations" are stressed by mathematical treatment.

7′. No significant difference.

8. Most of the topics in the course as a whole are presented without mathematics, and the course is entitled to the catalog description: "a nonmathematical approach to science."

8′. No significant difference.

V. Methods of Teaching

1. The students are asked to use one basic text, and it is selected from a list of five current books. Many supplementary texts are consulted throughout the year, and the college library contains a large number of reference books to be used in connection with the course. The instructor is not very well satisfied with the text he is using, believing it to be too "shallow" or

1′. Instructors here select their text from three titles, two of which are among the five used in *Science Survey* courses. They offer the same criticisms of existing texts.

"comprehensive," on the one hand, or too "technical" or "rigorous," on the other, his criticism depending upon the particular text he has chosen.

2. Individual laboratory work is not included, but considerable emphasis is placed on demonstrations as a substitute. The demonstrations are generally confined to topics which the instructor feels might be difficult for the students to understand. The students are asked to remember the principles that have been demonstrated, but are not required to prepare reports on the demonstrations. Two-thirds of the reporting instructors state they have no laboratory work.

2′. The same ratio reports no laboratory work. In its place, the instructor demonstrates each topic fully, and the students are asked to remember, but not to write reports on, the principles demonstrated.

3. Motion pictures, lantern slides and other visual aids are used to illustrate the topics; in colleges and universities, probably 25 films are shown each semester; in a junior college, a larger number. The students are expected to remember the principles illustrated, but no report is required.

3′. Fewer lantern slides and motion pictures are shown, but here, also, the students are asked to remember, but not to write on, the principles illustrated.

4. Field trips are a negligible factor.

4′. A field trip is very unusual.

5. The student is expected to spend 1.5 to 2 hr in out-of-class preparation. It is expected that most of this time will be spent on study of the text, and that reports on outside reading will take the next largest amount of time. Practically no time is to be spent on mathematical problems, individual investigations or projects.

5′. Same.

6. Standardized tests are not given, but the instructor feels a great need for them, and would like to give them, especially on subject matter, but also on attitudes and aptitude. He feels that his present tests measure facts and memory work only, while the other objectives of the course go untested. His testing program is somewhat disturbed because the survey students have had such a varied background of science experience, but his greatest concern is to find tests that will measure adequately the "major" objectives of survey courses. He believes that the lack of an adequate testing program constitutes one of the major weaknesses of survey courses.

6′. Same.

7. The student's final grade is based principally upon written examinations. The instructor uses very little of his own opinion of the student's work or progress in determining the final mark.

7′. Same.

VI. Teacher Preparation

1. The instructor has had academic preparation well above the average. He, doubtless, has a doctor's degree either in physics or in chemistry, has majored in one and minored in the other, or has minored in mathematics.

1′. Same, except that the major is more probably in physics.

2. In spite of his splendid academic training, the instructor feels that it is not adequate for teaching this survey course, and recommends a broader type of training for teachers in this field. He particularly would like to see more work given in astronomy, chemistry, geology and geography, along with a broader training in the field of physics.

2′. This instructor also feels that his preparation should have been broader, but he does not feel as keen a need for training in the various sciences as does the *science survey* teacher.

3. Although the instructor would like to see the training broadened, he does not want it to be given by education departments, but either by the science and education departments working together, or by the science departments alone.

3′. Same.

4. The teacher wants training in specific skills which he feels would be useful for these courses. First of all, he would like to have more training in giving demonstrations; next, in testing "attitudes and appreciations."

4′. Same.

Certain teaching and administrative problems characteristic of these two types of courses became evident throughout the study, but limitations of space forbid their consideration here. From the evidence presented, however, it seems that the differences between the two courses are in few cases critical, the nearest approach to this probably being in the problem of whether or not the students should be "exposed," so to speak, to all the sciences, or "saturated" with one; and whether it is legimate, advisable, or worthy (depending upon the phraseology of particular

physics survey instructors) to utilize emotional appeal in propagandizing in behalf of science in these lay courses.

Some inconsistencies seem to become evident when certain course procedures are compared with the objectives held by the teachers. They say that the chief reason for adding these courses to the curriculum was the failure of regular academic and pre-professional science courses to meet the needs of the nonscience student. Yet, in their subject matter approach (with most of the mathematics deleted, it is true) they adhere amazingly close to the topics and stress of pre-professional courses. They wish to test the objectives of survey work other than facts and memory work, but, confessing that they know of no available tests for these other objectives and are themselves incompetent to construct them, they base the student's final grade entirely upon subject matter examinations. These are "layman" courses; yet such topics as "correct lighting," "electrical hazards," "home and office heating," "the automobile," "household appliances" and "the history of science" are almost entirely neglected, while certain theoretical topics get proportionately more attention.

The reasons for most of these problems seem to lie in the following situation. The instructors have convinced themselves of the purposes which these courses should serve, and have committed themselves to appropriate aims and objectives to realize these purposes; but, trained as they are in a highly specialized way for technical research, precision measurement and unemotional, objective approach to natural phenomena, they are unfamiliar with tools and machines outside their own research field, and they lack knowledge of the technics of testing and of where existing tests may be obtained. They are, therefore, as they themselves admit, in need of a broader training in all fields, and in special technics in this "survey" field, if they are to teach these courses adequately.

This study is presented as an entirely empirical determination of the existing standards and practices in the two types of courses. These standards have not been offered as "best," "good" or even "bad;" but sufficient time has elapsed since the courses were inaugurated in the curriculum, and a sufficient number of them are now in existence and have been canvassed in this study, to present them as "accepted" standards. It is hoped that the study will stimulate science teachers and their organizations to proceed to valid experimental work upon evaluation of survey course standards and the problems that have become evident as a result of their inspection.

The Teaching of Photography

Wilson Woodcock
Department of Physics and Astronomy, Hunter College, New York, New York

PROFESSOR Martin's survey[1] of the problems involved in organizing a course in photography and the solutions he has offered will be found invaluable by anyone contemplating the addition of such a course to the physics department curriculum. This note is an attempt to emphasize the uniqueness of the pedagogic problem inherent in photography. Martin touched on this when he said: "It must be borne in mind that the student is interested in learning the principles of photography for their own sake, and not merely as physical principles." This is most certainly true in the course which we give, where the majority of the photography students come from outside the physics department. Any shifting of emphasis from the primary objective of the course—the making of better pictures—to the exposition of physical principles for their own sake is invariably met with a loss of interest and decrease of activity on the part of the students.

From the standpoint of the student, this primary objective of the photography course may be interpreted in another way. The average student enters the course with an enthusiasm differing in kind from that evinced by the average

[1] Am. Phys. Teacher **7**, 116 (1939).

student in a straight physics course. In the latter case one finds an eagerness to comprehend a new body of knowledge, to understand the operation of machines, etc.; the desire is to grow in understanding. In the former case we are dealing with an immediate desire to create. Understanding of the basic principles of physics, chemistry and art is a means to an immediate end—the making of pictures; attention is focused on a principle because of its applications to a specific problem and not because of its importance in the structure of science. Where a large body of exact knowledge is necessary before creativeness may be attempted, as is the case with physics, it is to be expected that the rigor required for mastery of the science will be reflected in the teaching of that science. It does not follow that the same teaching technics that are successful in physics will be successful in directing students toward the mastery of a skill. To delay the student's taste of creation while a thorough understanding of physical and chemical principles is being imparted is to stifle most of the enthusiasm brought to the course. Rather, the enthusiasm should be nurtured and conserved to act as the driving force to achievement, both during the course and after.

The idea of catering to the student's urge to creativeness is by no means new. Dow's use of it in teaching sketching was a complete inversion of the conventional method.[2] His method begins with the study of composition and design, and proceeds to the refinement of technic in the drawing of details. In the older methods the student learned to draw *parts* of pictures before he was allowed a taste of the creativeness involved in putting the parts together into a picture. A similar change has taken place in the teaching of piano where, instead of beginning with scales and finger exercises as was once the custom, the student begins almost at once to play simple arrangements of the music he wishes to re-create.

Some of the ideas underlying the pedagogy used in the foregoing areas should be examined before the method is adapted to the teaching of photography. The first idea has already been stated, namely, that the course should be designed to utilize and cultivate the student's enthusiasm in the most efficient manner possible. This involves a further idea, that the effectiveness of a course is measured not entirely by the total number of facts learned by the student but also by the extent to which the student is inspired to learn for himself and is shown how to learn for himself after the course is completed. If the student acquires a great many facts and, at the same time, loses his taste for learning or gets the mistaken notion that he can learn only with the aid of a formal syllabus and a laboratory full of precision instruments, then the course is woefully ineffective. If he continues to learn on his own, as do many amateurs in photography, and finds that he is progressing more rapidly for having had the course, then the course is effective in my understanding of the term. Nor is it necessary to make the course easy in order to keep interest alive. It is essential, however, that the student be assured that the material he is learning is proper to his needs. If he finds it impossible to advance beyond a certain point in the making of fine pictures without a specified bit of information or understanding, or if he finds that a certain technic must be mastered before a desired result can be obtained, he is far more apt to respond favorably than if he is asked to take the assignment on the word of the teacher that the material is part of what every good photographer should know.

In the preface to his book, *The Enjoyment of Laughter*,[3] Max Eastman describes the method he employs to present his theory of humor. In the first chapter the theory is set forth as a whole. In the second chapter it is again given as an entity but with certain phases emphasized and refined. And so on to the end of the book the attempt is made to refine the reader's understanding without allowing him to lose sight of the central theme. Such a method, I believe, is peculiarly applicable to the study of photography. The many excellent self-taught amateurs use the method naturally and with success. Their knowledge of the theoretical background of photography advances more or less evenly on all fronts, and the end of each experiment is the making of a commendable picture. The achievement of a good photograph requires that each step in the photographic process be handled

[2] A. W. Dow, *Composition* (Doubleday, Doran, 1931).

[3] (Simon and Schuster, 1936).

satisfactorily *in the given situation.* Failure in any one phase of the process means failure to get a desired picture and provides incentive to eliminate the cause of that failure. A beautiful landscape spoiled by lack of depth of definition is a more natural spur to the understanding of the nature of depth of definition than is a set laboratory experiment which requires the same amount of effort and materials but which completely loses sight of the primary purpose of improving the quality of pictures. In like manner, the recitation program designed to help the student advance uniformly in all aspects of picture making will be found more congenial to the average student than a program that makes it necessary for him to become proficient in optics, then in the chemistry of the process, then in orthochromatics, and, finally, in the elements of composition.

A suggested order of topics for 15 recitation periods follows. The topics are essentially the same as those suggested by Martin; it is only the order of discussion that is different.

1. Demonstration and discussion of the photographic process from focusing of image to fixation of film; printing.
2. Optics: image formation with pinhole and with lens; viewfinding; focusing.
3. Exposure: problem of exposure; use of tables and meters.
4. Chemistry: composition of the photographic emulsion; chemistry of development and fixation.
5. Esthetics: elements of pictorial composition.
6. Orthochromatics: color sensitivity of films; use of filters.

Thus, in six sessions the student has seen the entire process demonstrated and has been introduced to the principal areas comprising the subject matter of photography. Meanwhile, through pursuit of the projects to be described later, he has been acquiring familiarity with all these areas in the laboratory. Some of the first pictures completed by the students will afford numerous examples of failures due to poor technic, lack of knowledge, lack of understanding of material already studied, etc. Such failures furnish the instructor with many excellent opportunities for reference to defects and remedies in connection with the second inspection of the subject matter of photography, which begins with the seventh session:

7. Optics: relative aperture and its relation to intensity of image; depth of definition and hyperfocal distance.
8. Exposure: more detailed discussion of exposure problems; use of the *f*: number system of aperture notation.
9. Chemistry: the developing solution: its constituents and the effect of formula variation on the characteristics of the developer.
10. Esthetics: elements of portrait lighting; lighting of still life.
11. Orthochromatics: further discussion of films and filters. Optics: defects of lenses; the photographic lens.

The remaining sessions are devoted to the study of emulsion characteristics, and to the discussion of "special processes" which require as prerequisite an understanding of the basic process. Color photography is suggested here because of the widespread interest in the subject and because it utilizes so much of the student's newly acquired knowledge.

12. Exposure: characteristic exposure-density curve of the photographic emulsion.
13. Exposure: relation of exposure and development to negative characteristics; characteristics of chloride and bromide papers.
14. Color photography: additive process.
15. Color photography: subtractive process.

We have not had experience with photography courses comprising 30 or more sessions and are therefore unable to say whether or not the method here described could be used successfully in an expanded, extended form. However, it is our opinion that the additional periods would make it possible to continue this procedure of expanding and refining the students' understanding of photography in an effort to approach the instructor's concept of a rigorous treatment of each phase of the subject.

Our laboratory work consists of a series of project assignments differing from the conventional experiments in that (1) the issue of each project is a picture or set of pictures of which the subjects are, for the most part, chosen by the student; (2) the printed "instructions" are not instructions at all but a page containing suggested references, difficulties to be avoided, a few hints on fine points of technic and two or three suggestions on how to improve the photographs as pictures; (3) some of the work is done at home and some in the laboratory at other than the scheduled hour, thus freeing part of the laboratory time for demonstrations.

Whenever it is necessary to control some of the factors in the photographic process while the student is occupied with others this method approaches the conventional

laboratory procedure. For example, to learn the technic of development the student may be instructed to use a roll of chrome-type film exposed in a prescribed manner; or, in the study of filters, the student may have acquired control over exposure, development, etc., by the time the consideration of filters is introduced.

Projects are preferred for a number of reasons. Martin has pointed out their advantage where darkroom space and apparatus are limited. They are also better suited to the triple purpose of sustaining interest, of affording practice in self-education and of keeping the emphasis on the avowed objective of the course, the making of better pictures. Some projects that in practice seem to satisfy these requirements are the following:

1. Construction and use of a pinhole camera.
2. Developing and printing technics.
3. Developing and printing at home.
4. Preparation and use of filters.
5. Projection printing.
6. Improvisation of projection-printing apparatus at home.
7. Portraiture.

The value of the project on the pinhole camera justifies the prominent place given it in many textbooks and courses of study. The student learns at once the necessity for a light-tight box and for keeping the camera rigid during the exposure. He learns how to manipulate cut film and how to improvise a viewfinder. The problem of exposure is simplified by the elimination of one variable, for the relative aperture is fixed. The variation in exposure time is a matter of seconds, which is much more real to a beginner than is a variation from 1/25 to 1/50 sec. The student is introduced to the exposure table and, in general, is quick to see its value. The pinhole camera is a real addition to the photographic equipment of the beginner, for there are many types of subjects for which it is more suitable than the cheaper lens cameras. Questions concerning all phases of the photographic process occur to the students as they work on this project, and valuable laboratory discussions can be built around them.

Another successful project is one in which the student is asked to develop and print pictures in an improvised darkroom at home. The routine followed in the college darkroom is sometimes accepted uncritically by the student until he attempts to set up his own routine with his own equipment. Then many questions arise. Is it necessary that the room be absolutely light-tight while developing panchromatic film? How much white light can be tolerated while making contact prints? Is a box type printer indispensable? Is a ferrotype plate for glossing prints indispensable? Is there a simple technic for flattening prints? What is the minimum time required for print and film washing? The student also sees the developing and printing routine used in the laboratory in a new light as a goal of efficiency to be equaled or surpassed in the home routine. Many of our students have been led to the construction of safelights, box printers, washing tanks, etc., with a gratifying amount of learning attending the process.

The list of projects given is not intended to occupy an entire semester's laboratory time, though it may be adequate for the slower students. As soon as a student has completed these projects—those who enter with some experience as amateur photographers are allowed to omit certain of them—he is allowed to choose additional projects from a list containing many of the topics suggested by Martin.

This note represents an effort to express a point of view—that a course in photography is simply that, and not a study of physical principles from the standpoint of their application to photographic practice. In the majority of cases the course is taken by students desirous of becoming proficient in a rapidly growing popular art. This is an incentive greatly different from that inspiring the physics student. We should reflect carefully on this difference of objectives before we hasten to transplant in entirety the pedagogic methods of physics to the photography course.

In order to arrive at the truth, we are dependent on an adequate supply of facts, and not only on accurate methods of thought.—Robert H. Thouless, *How to Think Straight* (1939).

An Experiment in Laboratory Instruction

ROBERT E. BERGER
Bethany College, Bethany, West Virginia

IN his book, *The Teaching of Physics,* C. R. Mann decries the lassitude which one usually finds in an elementary physics laboratory. He describes the intense feeling that possessed Galileo and Archimedes after they had made some of their first discoveries and asks, "Is there any teacher today anywhere who ever observed any boy or girl become enthused with any such emotions as these after performing any one of the 'forty experiments from the following list'? If any, speak; for him have I offended." I believe that the majority of those who teach physics will agree with Mann that the usual laboratory class is lacking in life and interest. One can find, without much difficulty, many articles concerning the possible use of a scientific method—the deductive method, the inductive method and the like—but little or nothing is available concerning the actual application. The four or five main steps which are usually given for these methods are rules to be practiced for several days and then conveniently forgotten.

With the permission of Dr. J. S. V. Allen I was able to apply, in my capacity of undergraduate assistant in the elementary physics laboratory of Bethany College, a method of instruction which I believe has some possibilities for the development of a truly scientific attitude in the majority of students. This method was applied for a period of six weeks to 15 students divided into two groups. The results are difficult to estimate because of the lack of control groups, achievement tests and the proper amount of time; hence I present this material not as definite results but in the hope that others may be interested in applying it.

As a beginning, the laboratory manual was dispensed with entirely. When the student entered the laboratory he was assigned to an individual table upon which were placed certain pieces of apparatus used in the study of some general topic, such as sound, magnetism or electrostatics. The rotation system was used so that each student worked at each table but on different days, and two hours was the minimum time spent working at each table. The only rules were that the student keep busy and that he make a detailed report of his activities. The report was written while the work was being done, the following general outline being used:

1. Description of apparatus.
2. Observation and investigation of phenomena.
3. Conclusions reached concerning phenomena.
4. Check at end of the period in textbooks to determine if conclusions were correct and to find the cause of error if they were wrong.

The grade received by the student at the conclusion of the six weeks' period was based primarily on two factors: the manner in which he conducted his investigations and the manner in which he described his activities and stated his conclusions.

The choice of apparatus for each of the tables was the most difficult task. When selecting and arranging each piece I endeavored to keep in mind the idea that I was trying to sell it to the student. In short, I was attempting to arouse in the student's mind a natural curiosity and interest. At the table dealing with sound were placed tuning forks of various frequencies, open and closed organ pipes of various lengths, a small xylophone and a simple sonometer. The table devoted to practical electricity contained a wide assortment of electrical switches, a doorbell, fuses, household appliances, ammeter, voltmeter, batteries, and a source of 110-v direct current. A complete list of apparatus to be used, together with notes on the necessary precautions, was given to the students at the beginning of the six weeks' period. The care in the use of ammeters, voltmeters and electroscopes was especially stressed.

At the close of each period I took notes concerning some of the unusual things which I had observed and upon them I base the following comments and conclusions:

From the first day I noticed that almost invariably the students who had previously had physics in high school accomplished less than those who were without this experience. The only possible reason I can give is that the students who had already taken physics spent most of their time trying to remember and repeat their simple high

school experiments. This type of student presents a serious problem to anyone undertaking to apply this method of teaching.

During the first few weeks the majority of the students observed many phenomena without realizing that each one was something which could be investigated further. One good example was noted at the "magnetism table." One of the several bar magnets placed there had a north pole at both ends. It was observed in two different periods that both poles repelled the north pole of a compass before one of the students decided to check and investigate the problem. This student spent most of the period determining by experimentation the possible causes of the discrepancy.

The table which was most popular throughout the six weeks was the one concerned with "practical electricity." Nearly everyone began by trying to find the differences between the various types of electrical switches. Many had never seen the inside of even the simplest switch. Parallel and series circuits as used in lighting and doorbell systems also interested many. One girl decided to estimate the cost of the electricity she used in preparing midnight lunches, and was surprised at the low cost. Few students seemed interested in the fuses which were available for study.

At the "sound table" one student devised an interesting experiment showing some originality. In the hall outside of the laboratory there is a large ventilating fan which runs at various speeds. By means of tuning forks and the sonometer he studied the frequencies of the sounds coming from the fan at these various speeds. He also produced a good demonstration of the phenomen of beats.

Occasionally something misleading may be observed by the student and such things must be carefully explained. For example, a girl at the table arranged for the study of the magnetic effect of a current asked me to suggest something for her to study, so I told her to twist the wire into a coil after she had investigated the effect with a straight wire. About an hour later she told me that she had magnetized a stick of wood. It seems that she had placed an iron rod in the coil, had tested it for polarity in this position, and then had substituted the wooden stick for the iron rod. The test for polarity gave the same result and from this she had reached her conclusion. If she does not remember anything else from this experience she should at least know that nothing is proved by a single test. It was interesting to note that several students, while studying the magnetic effect of a current, constructed simple galvanometers after they had seen one used in a class demonstration.

The table devoted to electrostatics seemed to afford the least possibilities for investigation and, if I noticed that the student was inactive, I usually suggested that he try to find ways to give an electroscope a known charge. Several students who attempted to find how electrostatics and magnetism are related were directed into something more simple. One student in checking his conclusions at the end of the period found that his method of charging an electroscope by induction was exactly opposite to that described in the book. He decided to remain and investigate further, and found that the cause lay in the thick crepe soles of his shoes.

One table was equipped for a study of simple wave motion, but because this subject can become very complicated I always worked with the student and directed the experiment more closely.

For students who were especially proficient in some special field of physics the procedure was varied considerably. For example, one student was interested in sound, and when the time came for him to attack this subject I asked him to experiment with the Kundt tube. In two weeks he was able to give a very good demonstration of this phenomenon, having begun with only the glass tube, metal rods and a textbook description. Another student who had asked for something more difficult in electricity produced a good potentiometer set-up. Such instances were few, however; the majority of the class followed the original plan.

Some objections have been voiced concerning this plan. One experienced teacher who visited the laboratory believed that the students should not work so blindly. He suggested that more definite problems be assigned to them and cited, as one example, the comparison of the amounts of moisture which leave the body through the lungs and through the skin. This problem seems to me to be more suitable as an advanced experiment than for the elementary laboratory. As for the blind methods of the students, my only reply is that any individual who has been told for the first twelve years of his educational experience exactly what to do and how to do it cannot be expected to learn in a six weeks' period how to depend on his own eyes and brain.

An objection raised by some of the students was that few definite numerical results were obtained. These students should be given sufficient opportunity to do quantitative work in the advanced classes, but it seemed best to place the emphasis on qualitative work in the elementary laboratory.

It seems to me that I have observed the following things among the students in this small group. The student (1) learns to discover and observe for himself phenomena for possible further study, (2) learns to develop his own methods of investigation, and (3) develops skill in describing his activities and in stating his conclusions in a clear and comprehensible manner.

For the instructor I can guarantee from experience the busiest and most interesting periods of instruction he has ever supervised.

The Teacher's Motivating Interest

G. W. Stewart
State University of Iowa, Iowa City, Iowa

SCIENCES, and physics especially, have been charged with being merely collections of organized facts. The criticism has been made that science teachers treat these facts as the real essence of their subjects and consequently that there is little of cultural value—little of education, in the liberal arts sense—in them.

Naturally, we in physics are at first inclined to defend ourselves and perhaps to ask if there are any subjects in the liberal arts curriculums that are not guilty of the same fault. What subjects are taught without relatively too much emphasis upon fact? In what subjects is there an ample stress on those underlying aspects that are the source of creative interest? Are not the educational possibilities in many subjects too frequently unrealized?

But the act of defense may too easily prevent us from seeing our own faults clearly. We may compare two very different aspects met in the teaching of physics. One is that of phenomena or the performance in the physical world, and the other is the effort man makes to organize data, to interpret them and to express his understanding in mathematical form. The second aspect may really be called man's performance, man's very remarkable attempt to understand clearly.

The fact is, the amount of time consumed in this effort to understand far surpasses that involved in our direct contemplation of the phenomena. The same may be said of our teaching. We must necessarily spend almost all the time in the classroom and laboratory upon man's endeavor to establish "principles" from data. This time-emphasis on man's performance need not, however, dull our deep abiding interest in the phenomena of the physical world. Phenomena are the origin of man's inquiries and are the source of his inspiration. Toward the phenomena themselves every creative physicist looks. Doubtless, the productivity of researches in physics will always depend, in a very significant measure, upon the investigator's primary interest in the physical phenomena. We teachers of physics might reasonably ask whether or not we thoughtlessly allow our attention to be too much diverted from the performance in the physical world, from the "go" of things. Wherein do we test the interest and appreciation of our students in this performance? What is the influence of our examinations, of our class discussions?

I would not becloud this brief presentation by any discussion of method. The implied questions go very much deeper. Are we consciously motivated by our deep interest in what actually happens in the physical world? Or are we completely captivated with the marvel of man's effort to understand, thus being brought to regard the latter as the central thing? Is physics in a book or is the book a means to an end?

The object of this brief note is to emphasize this one aspect of teaching, and in so doing not to stress less our interest in man's marvelous achievements, but rather to stress more an increased conscious interest in the phenomena in the physical world. Indeed, an appreciation of man's effort is really essential to our appreciation of the phenomena itself, and of this we must not lose sight. Yet one of the glories of physics is that it is based on fact as much as upon sound reasoning. But the sincerity and depth of the more fundamental interest in the phenomena in the physical world is not to be determined by the amount of time spent on it in the classroom. This time may be relatively small, but the intensity and sincerity of interest is of very great importance.

What we need is not the will to believe, but the wish to find out, which is the exact opposite.—Bertrand Russell

Undergraduate Experiments for Determining the Boltzmann Constant and the Loschmidt Number

WALTER C. MICHELS AND SELMA BLAZER BRODY
Department of Physics, Bryn Mawr College, Bryn Mawr, Pennsylvania

DESPITE the accuracy with which comparatively inexperienced observers may determine the Boltzmann constant and the Loschmidt number by means of Millikan's oil drop experiment for e and a coulometer determination of the faraday, there are certain distinct benefits for the student in a direct kinetic theory evaluation of these important constants. The set of experiments to be described below has been used for several years in a second-year course in thermodynamics and kinetic theory, and has served to give the students a related laboratory program during a one-semester course. No attempt has been made to develop precision apparatus or methods, but the technics used are sufficiently novel for students at this level so that their interest is captured and they are encouraged to exercise their abilities. The accuracy is sufficient to match fairly well the first-order theory taught in the lectures. Some historic interest is aroused by the similarity of the methods used to the original work of Loschmidt.[1]

Because of the convenient location of its critical point, carbon dioxide is an almost ideal working substance for the experiments. It has been found that the commercially available gas is sufficiently pure (99+ percent) for the accuracy of this work and thus problems of contamination, difficult to solve without rather elaborate precautions with laboratory generated gas, are avoided. For a determination of the density of solid carbon dioxide, commercial dry ice is used.

Preliminary tests with a Boyle law apparatus of conventional form indicate that carbon dioxide at room temperature and atmospheric pressure is described by the ideal gas law to within a few percent. Consequently, the determination of the gas constant R is made by filling a glass bulb of about 1 l capacity with carbon dioxide at measured temperature and pressure, and weighing on an analytic balance. Weighing next with the bulb evacuated and then with it filled with water, one obtains the tare mass and the volume. Then R may be found from the ideal gas law, assuming that the gram molecular weight of carbon dioxide may be taken as 44 gm from chemical evidence.

With R known and spherical molecules assumed, the evaluation of the fundamental constants of kinetic theory reduces to the problem of finding the molecular radius σ. The gram molecular volume, $4\pi N\sigma^3/3$, may be determined from a study of the isotherms of the gas-liquid system at high pressures or, with the help of an x-ray study or of reasonable assumptions, from the density of the solid at atmospheric pressure. We employ both methods and make use of the agreement between them as a check on the validity of the assumptions. The gram molecular cross section $4\pi N\sigma^2$ may be found from the analysis and measurement of any transport problem. Measurements of thermal conductivity and diffusion have been found to involve an experimental technic somewhat too advanced for students at this level and we have consequently settled upon the viscosity as being the most suitable property for this determination. From the gram molecular cross section and volume one may determine N and σ.

The apparatus used for studying the isotherms is shown in Fig. 1. A hand pump produces pressures up to 300 kg/cm² on cylinder oil. This pressure is transmitted by means of a mercury piston to the gas, which is contained in a Pyrex capillary tube of inside diameter 3 mm and

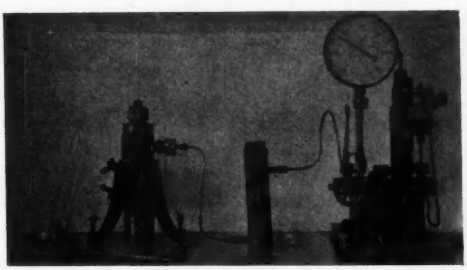


FIG. 1. Isotherm apparatus for carbon dioxide.

[1] Wien. Ber. (2a) 52, 395 (1865). See also E. Meyer, *The Kinetic Theory of Gases*, tr. by R. E. Baynes (Longmans, Green, 1899), Chap. 10.

outside diameter 10 mm, shown in cross section in Fig. 2. The upper 15 cm of this capillary is surrounded by an open water jacket in which the temperature is observed and maintained constant to about 0.2°C by stirring and the addition of warm or cold water. With temperature and pressure held at desired values the heights of the top of the mercury column with respect to the top of the tube can be obtained with a cathetometer. These height readings are then converted to volumes by use of the assumption that the tube has cylindrical symmetry. The apparent radius of the bore is measured with a filar eyepiece and corrected for the refraction of the glass[2] by dividing by 1.515. Isotherms showing liquefaction are taken at 4°C and room temperature, and a third run, to show the behavior above the critical point, is ordinarily made at 35°–40°C. At the lowest pressure which brings the mercury piston into the visible range (about 40 atmos), and at this highest temperature, carbon dioxide is sufficiently well approximated by the ideal gas law so that n, the number of moles of gas present in the capillary, can be obtained with fair precision by neglecting the a and b of the van der Waals equation of state,

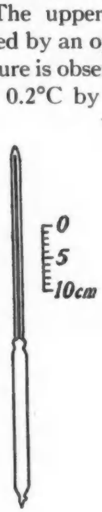


FIG. 2. Capillary compression chamber.

$$(p+a/n^2V^2)(V-nb)=nRT.$$

The isotherms are then plotted as V *versus* $1/p$ curves and extrapolated to infinite pressure. Since the volume at this value is, to the accuracy of the van der Waals equation, $nb=16\pi N\sigma^3/3$ for all temperatures, a direct determination of the gram molecular volume is obtained.

The density determination for the gram molecular volume is carried out by a flotation method. For the two liquids, respectively denser and less dense than solid carbon dioxide, trichloroethylene and acetone are used. Even when these liquids are precooled and saturated with carbon dioxide, evolution of gas causes some difficulty. However, if the behavior of a small piece of the solid is observed immediately after it strikes the mixed liquid in the Dewar flask, the point at which the densities of the solid and of the liquid mixture are the same may be judged with considerable precision. When this point has been obtained, the density of the liquid is found by the method of Archimedes, by using a precooled brass cylinder. Assuming a simple cubic lattice, as closely packed as possible in this form, one finds the molecular radius to be related to the density by the equation

$$8\sigma^3N=m/\rho,$$

where m is the gram molecular weight and ρ is the observed density.

A simple apparatus for the determination of the viscosity is shown in Fig. 3. It consists of two Pyrex bulbs, each of about 1 l capacity, joined by a 1-mm capillary tube about 50 cm long. To each bulb is attached a closed-end manometer and a stopcock for connection to a tank of carbon dioxide or a vacuum pump. An initial pressure difference is obtained by pumping one bulb while gas is admitted to the other. After the two stopcocks are closed, the rate of return to equilibrium may be observed by reading one manometer at 5-sec intervals for about 1 min.

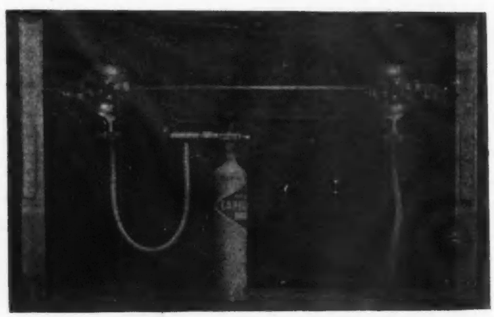


FIG. 3. Viscosity apparatus.

[2] Michels, Am. Phys. Teacher 7, 258 (1939).

Using the Poiseulle formula,[3] one can show the time dependence to be given by the equation

$$\log\frac{p_2-p_1}{p_1}=-\frac{\pi a^4(p_2+p_1)}{8L\eta V}t+\text{constant},$$

where p_2 and p_1 are the pressures in the two bulbs, a is the radius of the capillary, L is its length, V is the volume of a single bulb, η is the coefficient of viscosity of the gas and t is the elapsed time. A set of readings from only one manometer is sufficient because equal volumes of the two bulbs lead to the condition, $p_1+p_2=$ constant. The coefficient of viscosity is obtained from the straight line plot of $\log[(p_2-p_1)/p_1]$ *versus* t, and from it the gram molecular cross section is calculated by the usual formulas.

These determinations of $N\sigma^3$ and $N\sigma^2$ provide for an immediate solution for σ, and hence for N and k. A typical set of data obtained by students is summarized below:

Gas constant R	$(9.0\pm0.4)\times10^7$ erg mol^{-1} deg^{-1}
Van der Waals constant b	69 ± 6 cm^3
Density of solid ρ	1.6 ± 0.1 gm cm^{-3}
Viscosity η (23°C)	$(1.7\pm0.4)\times10^{-4}$ dyne sec cm^{-2}
Gram molecular volume:	
From isotherms	17.2 ± 1.5 cm^3
From solid	14.4 ± 0.8 cm^3
Gram molecular cross section	$(2.89\pm0.70)\times10^9$ cm^2
Molecular radius σ	1.93 ± 0.45 A
Loschmidt number N	$(6.1\pm3.0)\times10^{23}$ mol^{-1}
Boltzmann constant k	$(1.47\pm0.75)\times10^{-16}$ erg deg^{-1}

[3] *Handbuch der Physik*, Vol. 7, p. 103.

It will be noticed from this summary that the greatest source of uncertainty is introduced by the error in the viscosity measurement. The simple apparatus shown seems to be reasonably free from systematic errors, and the large source of uncertainty occurs in the pressure readings on a moving manometer at short time intervals. We plan to avoid some of this uncertainty in the future by the use of electrical timing, which, it is hoped, will bring the determination of the cross section into line with the rest of the work and thus reduce the uncertainty in the final constants.

We have found that the carrying through of these experiments has given to students a sense of the reality of the kinetic theory concepts. The work has been particularly valuable for those who have been able in successive semesters to determine N and k by these and by electrical methods.

The Magnetron Experiment and the e/m Ratio for Electrons in an Intermediate Course

D. H. Tomboulian and P. L. Hartman
Department of Physics, Cornell University, Ithaca, New York

THE experiment to be described is intended to be of the group of experiments in electronics usually carried out in an intermediate course in physics. Most of these experiments deal primarily with the motion of ions in electric fields. However, it is of some value to present to the second-year student an experiment in which the influence of magnetic, as well as electric, fields is studied, provided the observations and calculations are sufficiently simple to carry out. The particular procedure and treatment to be outlined here was recently introduced in a sophomore course designed to meet the needs of students with majors in engineering and chemistry, and proved to be quite instructive. The students, in groups of two, performed the entire experiment, after the circuit had been inspected by the instructor.

The experiment is based upon that of A. W. Hull.[1] It includes the determination of the characteristic curves of a magnetically controlled diode and an evaluation of e/m. It employs standard equipment and such apparatus as can readily be constructed in a laboratory having average facilities. Fig. 1 is a diagram of the magnetron having cylindrical symmetry. The nonmagnetic anode A is a cylinder of diameter 1.91 cm and length 5.00 cm, made of 5-mil sheet tantalum. The cylinder was formed and spot-welded while on a mandril of known diameter. The filament F is a 5-mil tungsten

[1] Hull, Phys. Rev. 18, 31 (1921).

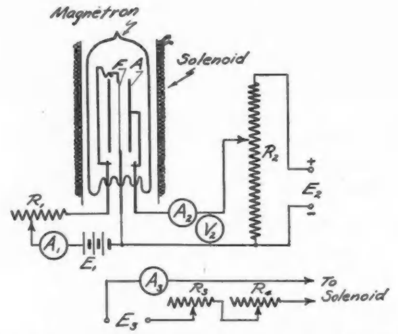


FIG. 1. Diagram of the magnetron and associated circuit. A_1, ammeter, 0–3 amp; A_2, ammeter, 0–100 ma; A_3, ammeter, 0–5 amp; V_2, voltmeter, 0–150 v; R_1, rheostat, 10 ohm, 5 amp; R_2, rheostat, 1000 ohm, 1 amp; R_3, R_4, rheostat, 23 ohm, 5 amp; E_1, 12 v; E_2, d.c. generator, 110–220 v; E_3, d.c. generator, 55 v.

wire of length a little less than that of the cylinder and was made to coincide with the axis of the anode by visual alignment. The tube, complete with getter, was constructed, evacuated and baked in our laboratory.[2]

The axial magnetic field is furnished by a two-layer solenoid having about 17 turns of No. 18 copper wire per centimeter. It is wound on a brass tube 8 cm in diameter and 40 cm in length. The ammeter A_3 and the voltmeter V_2 should be calibrated, as the precision of the results depends particularly on the accuracy of these two meters.

The following observations are made. With the filament current at 2.2 amp ($T = 2500°$K), the values of the plate current corresponding to a series of magnetizing currents are recorded for given potential differences between the filament and anode. This potential difference must be kept constant during a single run. A typical set of data for a range of voltages from 40 to 110 v is represented graphically in Fig. 2. A shown by the curves the cutoff region for the tube is quite sharp.

The steps in the evaluation of e/m by using the aforementioned data will now be discussed. Under the condition of cutoff, the relation between the plate potential V and the magnetic field H is given by[3]

$$V = \frac{1}{8}\frac{e}{m}b^2H^2\left(1 - \frac{a^2}{b^2}\right)^2 \text{ emu}, \qquad (1)$$

where e is the charge on the electron, m the mass of the electron, a the radius of the cathode and b the radius of the anode. If $a \ll b$, so that a/b is negligible relative to unity, the equation reduces to

$$V = \tfrac{1}{8}(e/m)b^2H^2. \qquad (2)$$

If the magnetic field is produced by a solenoid of length L and diameter D, the value of H on the axis of the solenoid at its center is given by[4]

$$H = 4\pi nIL/(L^2 + D^2)^{\frac{1}{2}} \qquad (3)$$

Substitution of (3) in (2) results in

$$V = k(e/m)I_c^2, \qquad (4)$$

where

$$k = 2\pi^2 n^2 b^2 L^2/(L^2 + D^2).$$

One factor which is important in determining the shape of the cutoff curves should be mentioned. The plate current, as seen in Fig. 2, does not drop abruptly to zero at a critical magnetic field. In using a directly heated cathode, of necessity there exists a potential drop along the length of the filament, so that if the plate voltage is V, measured at one end of the

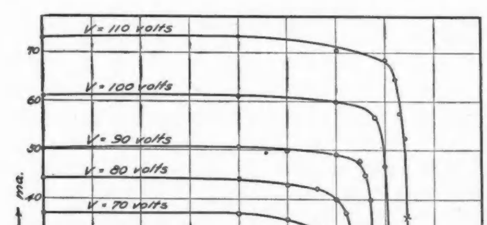


FIG. 2. Cutoff curves for the magnetron.

[2] It is gratifying to note that a new diode, FP 400, of approximately the above dimensions is now manufactured by the General Electric Co. This tube is designed for use in simple thermionic experiments in student laboratories and may also be used in the present experiment with reasonable success.

[3] For a proof of this relation see Harnwell and Livingood, *Experimental Atomic Physics* (McGraw-Hill, 1933), p. 119.

[4] Starting with the result found in Page and Adams, *Principles of Electricity* (Van Nostrand, 1931), p. 248, one can deduce this form for H by evaluating the angles in terms of the dimensions of the coil.

filament, it is $V+v$ at the other end, v being the drop across the filament. Then if the magnetic field is such as to prevent the flow of electrons to the plate from the low-potential end of the filament, electrons from the remaining portions of the filament will still arrive at the plate. Now, as the magnetic field is gradually increased, electrons originating at points of higher potential along the filament will successively be prevented

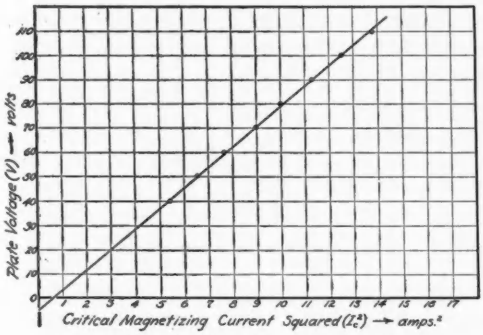


FIG. 3. Plot for the determination of e/m.

from arriving at the plate until finally electrons emitted at the point where the plate-filament potential difference is $V+v$ can no longer reach the plate. For this experiment the critical magnetic field is taken to be that which cuts off electrons from the mid-point of the filament, where the effective plate voltage is $V+\frac{1}{2}v$. Under this condition the plate current will have dropped to half its value with no magnetic field.

Equation (2) predicts that the plot of V as a function of I_c^2 should result in a straight line. The graph in Fig. 3 shows this plot for the same data plotted in Fig. 2. I_c is the critical magnetizing current and is chosen in the manner described in the preceding paragraphs. The magnitudes of the various quantities involved in the dimensionless constant k have the following values for this set of data: $n=16.8$ turns/cm; $b=0.953$ cm; $L=36.3$ cm; $D=8.10$ cm. Thus k is 4820. The slope of the line shown in Fig. 3 expressed in electromagnetic units is found to be 8.42×10^{10}. Hence, $e/m=8.42\times 10^{10}/4820=1.75\times 10^{7}$ emu.

In the evaluation of e/m using the slope of the aforementioned plot V *versus* I_c^2, the correction due to the contact difference of potential and the drop between the ends of the filament will appear as an additive constant to V, the voltmeter reading, and therefore will not influence the slope of the line. Indeed, if the foot of the cutoff curve, obtained by extrapolation of straight portion, be correlated with the voltmeter reading the result is also a straight line which has nearly the same slope as the line shown in Fig. 3 but a different V intercept. However, from the standpoint of the student, the method of interpolation discussed in this paper can be carried out with less ambiguity, since the vertical portion of the cutoff curve is not strictly linear but passes through an inflection point.

For the data treated in this paper, the intercept on the ordinate axis of the graph in Fig. 3 is 5 v. This should represent the constant correction due to the contact difference of potential and the potential drop across the filament. The measured drop across the filament for the heating current used is 7.5 v, which would indicate that an intercept of about 4 v is to be expected.

In an earlier publication[5] the use of a commercial radio tube was suggested. Practically all radio tubes which have cylindrical symmetry have additional electrodes. From the pedagogic point of view this complicates matters. Usually the collector in such tubes is nickel, which is ferromagnetic and distorts the applied field. The dimensions of electrodes are not accurately known nor are they usually of suitable magnitudes. Furthermore, the alignment of electrodes is not tolerably precise. The characteristic curves obtained from such tubes do not yield sharp cutoff curves and are not reproducible. On the other hand, the new diode manufactured by the General Electric Company[2] is very satisfactory so far as cutoff curves are concerned and can be used in this experiment satisfactorily, though they yield a value for e/m in error by as much as 10 percent.

Other recent articles[6] describe a variety of methods for determining e/m. In general these are suitable only for lecture demonstration or are designed for the more advanced student. The experiment here described can be carried out in the course of an hour and a half. The cutoff curves are quite sharp and the results are consistent. These are features that make the experiment worthy of inclusion in the intermediate curriculum.

The authors are indebted to Professor J. R. Collins, Department of Physics, Cornell University, for valuable suggestions during the development of this experiment.

[5] Hastings and Ohlgren, Am. Phys. Teacher **5**, 88 (1937).
[6] Bainbridge, Am. Phys. Teacher **6**, 35 (1938); Smyth and Curtis, Am. Phys. Teacher **6**, 158 (1938); Weber and McGee, Am. Phys. Teacher **7**, 62 (1939).

A Wave Machine and a Device for Compounding Two Simple Harmonic Motions*

H. W. Farwell
Pupin Physics Laboratories, Columbia University, New York, New York

I. A Wave Machine

SOME twenty years ago I devised an arrangement for showing various wave combinations, including stationary waves, but beyond building a crude model nothing was done; the idea was all right, but its realization was necessarily expensive, as probably all wave-model assemblies must be. The principles involved were not new and at least one other model has since been constructed along similar lines.[1] The desirability of demonstrating the various wave characteristics has not diminished, however, and recent reconsideration of the problem has resulted in the construction of a satisfactory wave machine with a reasonable expenditure of labor and money. Because this arrangement has certain distinct advantages, it may be of interest to present a brief description.

To indicate the sum of two displacements there appears to be nothing more clear than the case given by the linkage indicated in Fig. 1.

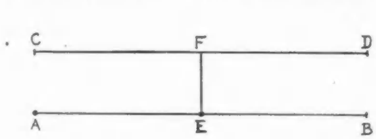


Fig. 1. Diagram of the linkage, constituting one element of the machine.

The rods AB and CD are linked at their centers by a tie rod EF. The end A is held on a fixed axis. The plane of the linkage is vertical, and when either C or D is given a vertical displacement, B will have the same displacement. If both C and D are given vertical displacements, B will have a displacement which is their algebraic sum. Such a linkage may then constitute the fundamental element of a wave machine.

So far there is nothing new, and the whole arrangement can easily be pictured as a matter of many parts, much adjustment and doubtful results. The two main difficulties are the construction of satisfactory elements at low cost, and the application of a workable wave form. The first difficulty has been solved by the use of standard steel umbrella ribs.[2] The joint was easily pushed to the center of the rib, the brace was cut 3 in. from the joint, and two of the ribs were then made into one element by soldering to the two braces a wire laid into the U of the cut ends. Thus were the 24 elements of the model quickly completed.

The hinge at A is simply a small copper tube soldered to a $\frac{1}{2}$-in. rod. In this tube there is milled transversely, each inch of its length, a slot which is just wide enough to engage the end A of the lower umbrella rib. For assembly the ends of the elements are inserted and a steel wire is run down the center of the tube.

The elements are now ready for mounting on a frame, which may be built with laboratory rods. To exhibit a static wave form all that need be added is a strip of plyboard or thin metal, which can be made in any form desired.

But the original plan was to exhibit combinations of moving waves, hence provision had to be made for some form of periodic displacement. The arrangement adopted was a rotor consisting

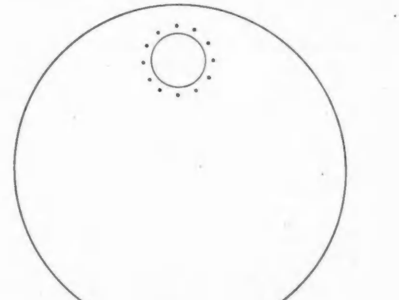


Fig. 2. Disk containing assembly hole.

* Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

[1] Everitt, J. Opt. Soc. Am. and Rev. Sci. Inst. **19**, 95 (1929).

[2] The ribs with braces were obtained from the Newark Umbrella Co., Newark, N. J.

of a $\frac{1}{2}$-in. rod carrying, at 1-in. intervals, circular disks $1\frac{1}{2}$ in. in radius, each with an assembly hole centered 1 in. from the center of the disk. These disks were obtained from sheet metal stampers. To aid in assembling, small holes were bored in each disk as indicated in Fig. 2. Since there are 12 holes, a phase difference of 30° between successive elements is readily obtained. As each disk is placed on the $\frac{1}{2}$-in. rod a suitable spacer is put next to it, and when all are on the rod a small-sized drill rod is threaded through that small hole in each disk which advances the disk 30° ahead of the previous one. The nut at the end of the rod is then tightened, thus clamping the disks firmly in place. Each disk would give a simple harmonic motion to the rib that it engages if the rib were infinitely long. With the ribs used the approximation is so close that any error thus introduced is less than the errors due to lack of straightness in the ribs.

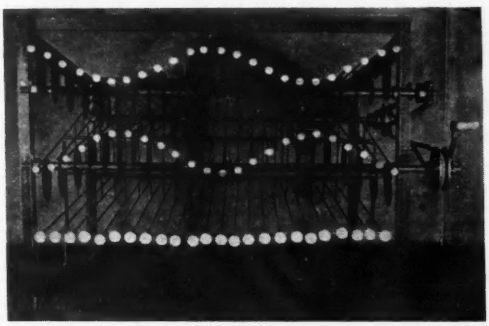


FIG. 3. Front view of the wave machine. The two components are opposite in phase.

The rotor can then be placed, for example, under the upper rib near D, with the edge of each disk engaging in the inverted U of the umbrella rib. When the rotor is turned a wave form travels along the D's, in a direction determined by the sense of the rotation. Thus waves may be made to traverse the C's and D's in the same or opposite directions, and the resultant wave will be shown on the B's.

As a luxury, yet highly desirable, the wave form on the C's can be brought approximately into the same vertical plane with the other two forms by attaching a third umbrella rib at C and running it at an angle to a point above D.

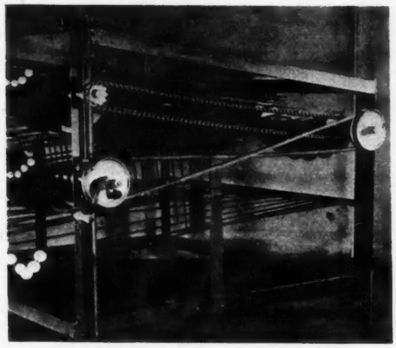


FIG. 4. Side view, showing the driving mechanism.

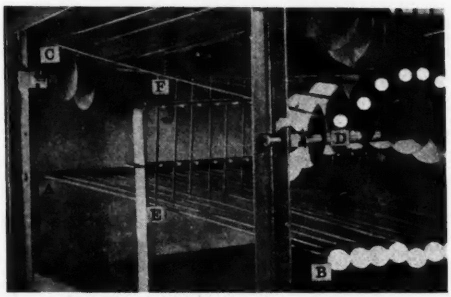


FIG. 5. Side view, showing details of the center guides.

A third rotor, identical with the one under C, and moving exactly in phase with it, will give this rib a motion of pure translation and its end, G, will present in proper plane the motion of C (Fig. 5).

There may then be shown in suitable fashion such combinations as these:

1. Single static wave forms.
2. Two static wave forms and their sum.
3. Two waves traveling in the same direction, with speeds alike or different, and their sum.
4. Two waves traveling in opposite directions, with speeds alike or different, and their sum.

Figures 3, 4 and 5 show the machine as assembled on a frame of angle iron.

II. A DEVICE FOR COMPOUNDING TWO SIMPLE HARMONIC MOTIONS

There are, of course, numerous methods for showing the sum of two simple harmonic motions,

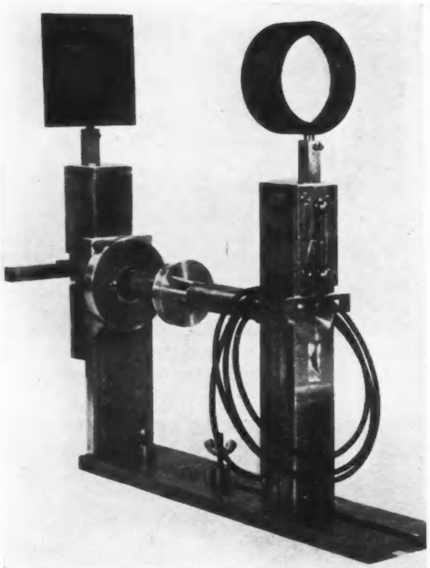


FIG. 6. Front view of the device for compounding two s.h.m.

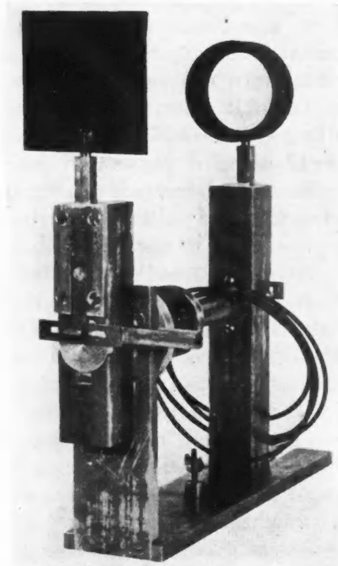


FIG. 7. Rear view of the device.

but no harm results from adding another, provided it possesses some advantages over those commonly in use. The arrangement here described will combine two such motions of the same amplitude and period, either when they are in the same direction or at right angles to each other, for any phase difference desired, changes in phase difference being quickly adjustable.

The apparatus (Figs. 6 and 7) consists of a suitable shaft mounted horizontally between two pillars; on each pillar is a slide containing a slot which engages a pin mounted eccentrically on a disk at the end of the shaft. Rotation of the shaft then imparts to each slide a vertical simple harmonic motion of amplitude 0.85 cm. On one slide is mounted a metal plate containing a hole about 0.6 mm in diameter, with a small opening, and on the other a lens of diameter 6.0 cm, and approximate focal length 22 cm. The distance between plate and lens can be adjusted so that when the apparatus is set in front of a projection lantern the lens will form on the lantern screen the image of the opening. The pillar carrying the slide which supports the plate can be rotated through 90° and the plate mounted to run horizontally instead of vertically; the lens moves only in a vertical path.

The shaft is jointed between the pillars, so that at the junction any phase difference may be obtained. The shaft also carries a pulley wheel and may be rotated by a hand wheel, since a slow rotation generally is desirable.

Because the images moves downward when the opening moves upward, it is the reverse of the opening's motion that is added to the lens motion; but this is easily understood and does not interfere with the comprehension of the phenomenon seen on the screen.

A college professor, asked by an illiterate inventor to criticize his design for a perpetual-motion machine, asked: "What about gravity?" To which the inventor replied: "Oh, t'Hell with gravity; we'll use plenty o' grease."—H. H. HIGBIE, *Lightning Calculations* (1934).

A Laboratory Experiment in Conical Refraction

J. R. Collins
Cornell University, Ithaca, New York

THE descriptions of experimental arrangements for the observation of conical refraction which one finds in textbooks of optics do not lead to a convenient or easily performed laboratory experiment. If, however, one views the virtual image of a very small hole formed by a plane-parallel plate of a biaxial crystal, the phenomenon is observed readily and, by means of a measuring microscope and a Nicol prism, the following information can be obtained: (a) the angle of incidence at which conical refraction occurs, (b) the angle of the cone of rays inside the crystal for the case of internal conical refraction, (c) the angle of the cone of rays outside the crystal for the case of external conical refraction and (d) the state of polarization of the light in these cones. For a crystal of known principal refractive indexes these quantities may be calculated by means of theoretically derived formulas and the theoretical results compared with those of experiment.

Figure 1 illustrates an arrangement used for this purpose. A brass tube T_2 was mounted on an axis perpendicular to the plane of the paper and a pointer attached so that the angular position of the tube could be read on a divided circle D. Another tube T_1 was fitted inside the tube T_2 so that T_1 could be rotated about its own axis. A biaxial crystal C was mounted in this inner tube with its plane-parallel faces perpendicular to the axis of the tube. This crystal was cut so that the principal axis, which bisects the acute angle between the optic axes, was perpendicular to the parallel faces. A sheet of thin metal foil covered the lower face of the crystal and a tiny hole was punched in the center.[1] For taking the pictures of Fig. 2 the hole used was about 0.5 mil in diameter.

A microscope with an objective of long focal length was mounted above the crystal so that the image of the hole was seen through the crystal. The hole was illuminated by focusing the filament of an incandescent lamp on the hole from below. The tube T_1 was adjusted by rotation until the plane containing the two optic axes was perpendicular to the axis of rotation and the microscope adjusted in position so that the line of sight was also in this plane. By rotation of the crystal, therefore, the light entering the microscope corresponded to any chosen angle of incidence from zero up to about 20°. When the angles of incidence were made 0°, 12°, 15.5° and 18°, the image of the hole appeared as illustrated in Figs. 2*A*, *B*, *C* and *D*, respectively. These pictures were taken by focusing the image on a photographic plate by means of the eyepiece of the microscope. The ordinary double refraction phenomenon was seen at all angles of incidence except near a critical angle at which two *ring images* appeared (Fig. 2*C*). The outer ring image was due to *internal* conical refraction, the case in which the *wave normal* was refracted along the *optic axis;* and the inner ring image was due to *external* conical refraction, the case in which the refracted *ray* was directed along the *ray axis*. Figs. 3(*a*) and (*b*) are ray diagrams for the particular crystal used (aragonite). The principal rays of the emergent beams were practically coincident for these two cases so that both ring images appeared at the same angle.

The outer ring image, of diameter d_i, was situated on the projected hollow cylinder of rays in Fig. 3(*a*), and its diameter was measured directly on the calibrated scale in the eyepiece

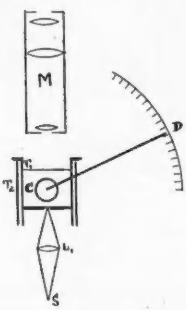


Fig. 1. Diagram of apparatus.

[1] This hole was made in the way described in Strong's *Procedures in Experimental Physics*, p. 69.

of the microscope. The expression for the angle θ of the internal cone of rays may be obtained as follows:

$$\theta=\frac{AC}{AO}=\frac{AB\cos r}{(t/\cos r)}=\frac{(d_i/\cos i)\cos r}{(t/\cos r)}$$

or $$\theta=d_i\cos^2 r/t\cos i. \quad (1)$$

The inner ring was situated on the projected rays of the external cone of Fig. 3(*b*). The angle ϕ of this cone is given by the expression

$$\phi=\frac{d_e}{AB}=\frac{d_e}{(t\cos^2 i/n_2\cos^3 r)}=\frac{n_2 d_e\cos^3 r}{t\cos^2 i}, \quad (2)$$

since AB is the distance from A to the virtual image of O formed by the parallel plate. The refractive index is the principal index n_2.

To obtain the angle of emergence i of these special rays forming the ring images, the two positions of the crystal were noted at which the ring images appeared on either side of the normal to the crystal face. Then i was half the angle between these two positions.

The calibration of the scale in the microscope was accomplished easily by viewing a tenth-millimeter scale as shown in Fig. 2*E*. The

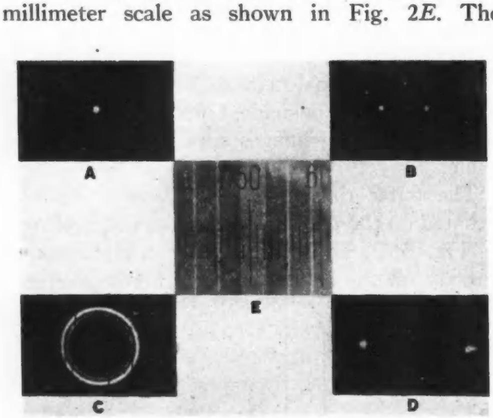


FIG. 2. Appearance of the image of the hole for various angles of incidence.

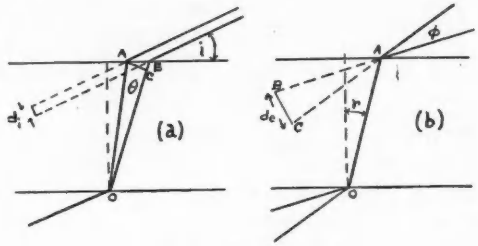


FIG. 3. Ray diagrams.

polarization[2] of the images was noted by placing a Nicol or Polaroid between the eye and the eye-piece of the microscope. At each point of the ring the light was found to be linearly polarized and with an angular difference of 90° between the azimuths of polarization at points diametrically opposite. Hence, the point of extinction on the ring rotated with twice the angular speed of the Nicol.

That each of the cones of rays is really a double cone[3] was seen by slightly defocusing: each of the rings broke up into two concentric rings with a dark ring between them. This structure, however, was not evident when the ring images were in good focus.

As examples of the numerical results obtained, the following measurements are quoted:

$d_i=10.4$ div., $d_e=8.8$ div.,
20 scale div. $=0.60$ mm,
$t=9.08$ mm, $i=15.5°$,
r (computed by using 1.681 for n_2) $=9°\ 10'$.

Equations (1) and (2) yield the experimental values, $\theta=2°\ 2'$, $\phi=2°\ 58'$. If the values of these angles are calculated by using the theoretical formulas given in the foregoing references and by using the principal refractive indexes of aragonite as given in the *Handbook of Chemistry and Physics*, the values obtained are: $i=15°\ 32'$ and $15°\ 27'$ for the inner and outer rings, respectively, and $\theta=1°\ 55'$, $\phi=3°\ 0'$.

[2] See Preston, *The Theory of Light* (ed. 5), p. 383.
[3] See Born, *Optik*, p. 240.

A scientific hypothesis must live dangerously or die of inanition. Science thrives on daring generalizations. There is nothing particularly scientific about excessive caution. Cautious explorers do not cross the Atlantic of truth.—LANCELOT HOGBEN, *Science for the Citizen* (1938).

The Cathode-Ray Oscillograph as an Aid to the Study of Some Electrical Principles

Herbert Trotter, Jr.
Department of Physics, Washington and Lee University, Lexington, Virginia

The advent of good, portable cathode-ray oscillographs has brought a new tool to the aid of physics teachers. The many problems to which this instrument can be applied to advantage in lecture demonstrations and laboratory work are just beginning to be realized.[1] This paper describes attempts that have been made to use this instrument in both lecture demonstrations and laboratory work as an aid to teaching some of the principles of electricity.

Charging and Discharging of a Condenser

Two of the electrical equations that a student can solve are those for the charging and discharging of a condenser through a resistance and through a self-inductance and a resistance. It is interesting to show the change of the charging or discharging curve from oscillatory to nonoscillatory. If the curves for the charging and discharging of a condenser are to be demonstrated and studied with an oscillograph, it is desirable that the traces of these curves should remain stationary on the screen of the oscillograph. This has been accomplished by repeating the trace on the screen at the rate of 60 times a second. Fig. 1 shows a diagram of the circuit by means of which this has been accomplished for the discharging of a condenser. The tube used is a type 885 Thyratron, a three-element gas-filled tube. A potential of 160 v was applied to the plate of this tube. Under this condition, as long as the grid is held at a negative potential exceeding 18 v with respect to the filament, no current will exist in the plate circuit. If this potential is kept on the plate and the negative potential of the grid is reduced to less than 18 v, electrons will flow to the plate and in so doing will ionize the argon gas in the tube. This ionized gas will now conduct a current of 0.3 amp. The grid cannot gain control of the current in the tube again until the current has stopped and the tube has had time to deionize. The negative bias is maintained on the grid by the potentiometer connected to the 22.5-v battery, and superimposed on this potential is a 60-cycle alternating potential of 5 v. Therefore, once during every cycle—that is, every 1/60 sec—the negative bias potential on the grid is reduced to less than -18 v. The Thyratron is thus used to charge the condenser 60 times a second, and the condenser under study, either C_3 or C_4, is then discharged through the desired inductance and resistance. The charge on the condenser C_1 or C_2 will leak off through its shunt resistance in time to increase the potential across the Thyratron so that the latter will recharge the condensers on every cycle of the 60-cycle alternating potential applied to its grid. The condensers in series with those under study allow 160 v to be applied to the Thyratron without having to apply this potential to the oscillograph.

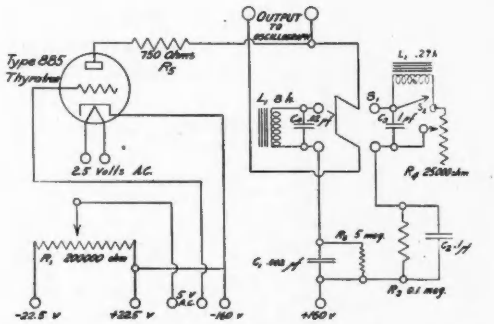


Fig. 1. Circuit illustrating discharging of a condenser.

The theory for the discharging of a condenser through an inductance and a resistance states that the discharging will be oscillatory when $R^2/4L^2$ is less than $1/LC$; that is, $R^2<4L/C$. If $R^2>4L/C$, the discharging will be nonoscillatory.

The electron beam of the oscillograph used in these experiments always travels from the right

[1] For examples, see the following articles in this journal: W. C. Bosch, **4**, 81 (1936); E. H. Green, **5**, 181 (1937); H. D. Smyth and C. W. Curtis, **6**, 158 (1938); A. H. Weber and J. F. McGee, **7**, 62 (1939).

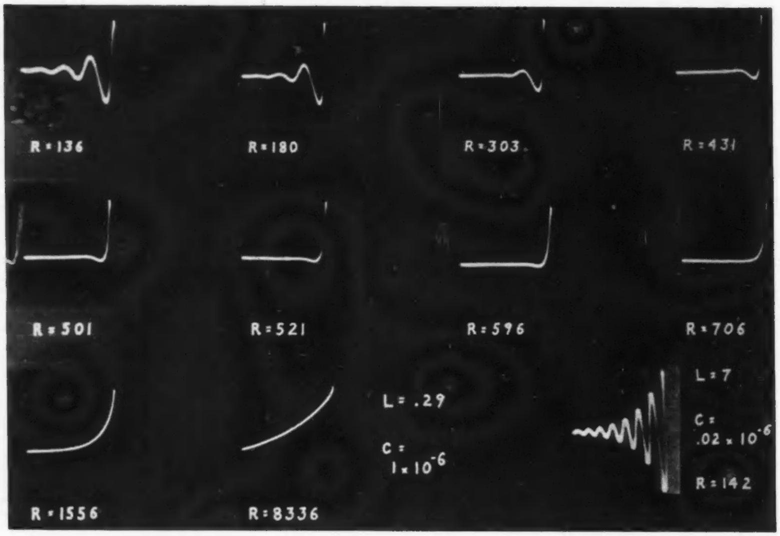


FIG. 2. Discharge of a condenser through an inductance.

side of the screen to the left. This must be borne in mind when the oscillograms are studied.

Figure 2 shows the oscillograms of the discharge of the condenser through an inductance. The capacitance of the condenser C_3, used to make ten of the oscillograms, was found to be 1.9 μf, and the inductance L_2 was 0.28 h. From these values the change from oscillatory to nonoscillatory discharge would be expected to occur at about 788 ohms. The oscillograms show a value slightly larger than 700 ohms. The last curve in Fig. 2 shows the discharge of a 0.02-μf condenser through an inductance of 8 h and a resistance of 142 ohms. The circuit is provided by reversing switch S_1 in Fig. 1. This oscillatory discharge has proved very helpful in explaining the operation of a Tesla coil.

Figure 3 shows the oscillograms for the discharge of the condenser through a noninductive resistance. Since the time required for the complete trace of the curve was 1/60 sec, the time required for the potential to fall to $1/e$ of its original value can be measured from these curves. The student can compare this value with the theoretical value, $t=RC$.

Figure 4 represents a similar circuit that can be used to study the charging of a condenser.

Phase Relations in A.C. Circuits

Figure 5 represents a circuit used to illustrate the phase relation of voltage and current in an alternating-current circuit. It consists of a transformer T_1 which can furnish a voltage in phase with the voltage applied to the circuit. The

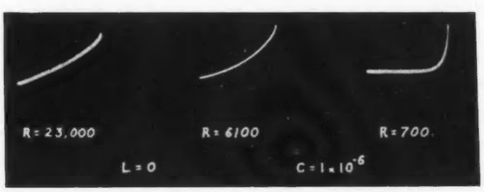


FIG. 3. Discharge of a condenser through a noninductive resistance.

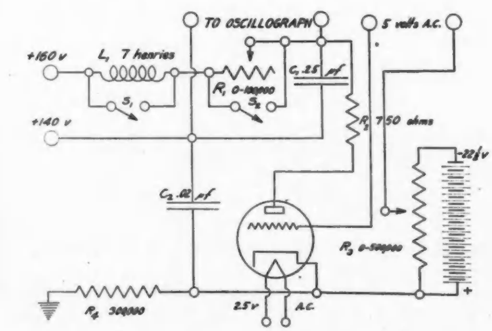


FIG. 4. Circuit illustrating charging of a condenser.

second circuit consists of a high inductance, a variable bank of condensers, a variable resistance and a fixed noninductive resistance. Output *B*, the potential across the noninductive resistance, which is always in phase with the current through the circuit, is applied to the oscillograph. The phase relation of output *A* and output *B* is

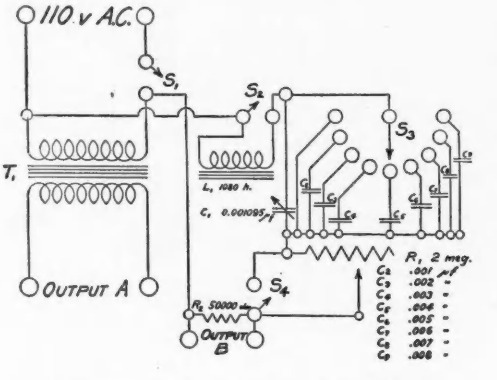


FIG. 5. Circuit illustrating a.c. phase relations.

studied by the oscillograph. This is accomplished by the use of an electron switch[2] which applies both outputs *A* and *B* to the oscillograph at the same time. Both of these outputs are made to appear simultaneously on the screen by having the electron switch apply first one and then the other to the oscillograph. The switching from one output to the other occurs about 60 times a second, so the two patterns appear stationary on the oscillograph screen.

Figure 6 shows oscillograms that have been made with this arrangement. The first one shows the current and voltage in phase when there is only resistance in the circuit. With inductance and resistance in the circuit the current is seen to lag the voltage in the second curve. The third curve shows that the addition of more resistance decreases the angle of lag of the current. The fourth shows that capacitance and resistance give a leading current. As the resistance is increased in the fifth and sixth curves, the angle of lead is decreased. The seventh shows the case of inductance, resistance and capacitance in the circuit with the reactance of the capacitance greater than that of the inductance and a leading current thereby resulting. As the capacitance is increased, its reactance is decreased and the eighth curve shows the current lagging the voltage. The ninth curve shows the adjustment of the capacitance to secure a resonance circuit with the current and voltage in phase. Since the student can see the shift in phase as *L*, *C* and *R*

[2] The electron switch was manufactured by Allen B. Du Mont Laboratories, Inc., Passaic, N. J.

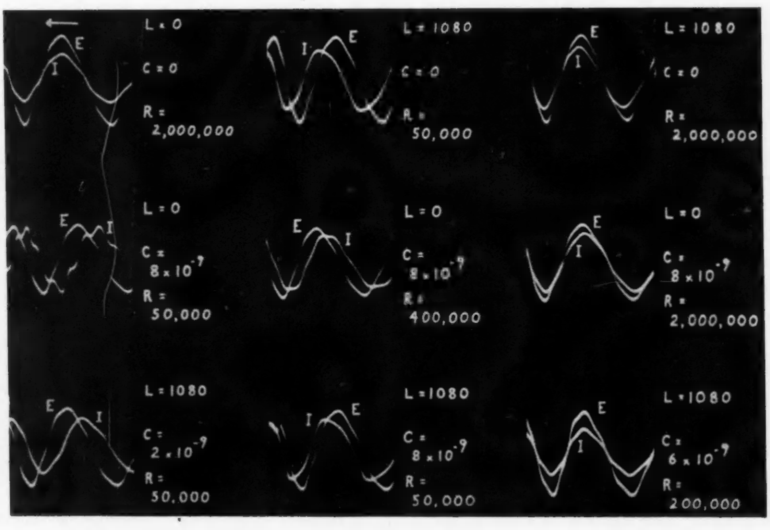


FIG. 6. Phase relations in an a.c. circuit.

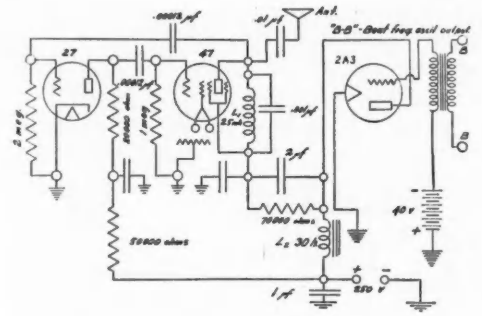


FIG. 7. Circuit of modulated oscillator.

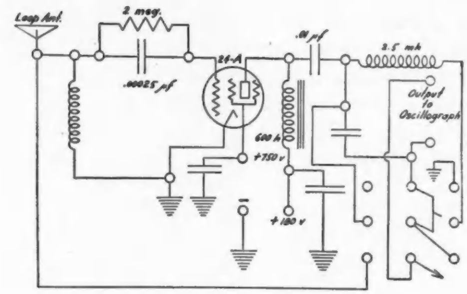


FIG. 8. Circuit of radio receiver.

are varied, this apparatus has been found to be very helpful in teaching alternating-current theory.

FUNDAMENTALS OF RADIO COMMUNICATION

In attempting to teach beginners the fundamental principles of radio communication, the instructor generally shows the experiment of lighting an electric lamp in a tuned receiver by the energy received from a high frequency radio oscillator placed some distance away. This experiment is useful for showing that energy can be transmitted through space, but aside from this it does not demonstrate the principles of radio communication. To accomplish the latter, a small radio oscillator and a receiver have been constructed and their operation shown on the oscillograph. Fig. 7 is a diagram of the modulated oscillator. The circuit containing tubes 27 and 47 acts as a radio oscillator with a frequency of about 10^6 cycle/sec. This carrier frequency can be modulated by a beat frequency oscillator connected to the terminals *B–B*. By using a constant frequency for modulation, the pattern can be held stationary on the screen. This, of course, could not be accomplished if a varying sound pattern were employed.

Figure 8 is the diagram of the receiver. The switching arrangement allows the oscillograph to be connected directly to the input of the receiver, or to the output of the detector, or to this output after it has been sent through an inductance. The oscillograms shown in Fig. 9 were obtained with this circuit. The first curve shows the carrier wave when the electron beam is swept at a rate of 10^5 times a second across the oscillograph screen. The second curve shows the modulated carrier wave and the third, the output from the detector.

Figure 10 is a photograph of the radio demonstration apparatus. The student can see that the signals are being picked up from one antenna to the other by noting the change in the height of

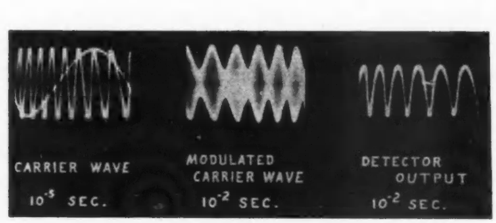


FIG. 9. Radio wave forms.

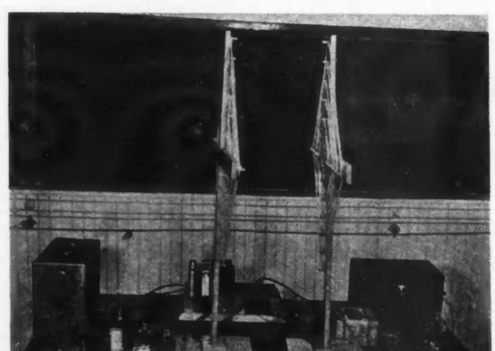


FIG. 10. A radio demonstration apparatus.

the wave on the oscillograph as the distance between the antennas is varied.

A Convenient Power Pack

The power pack shown in Fig. 11 has been constructed to supply all the necessary potentials for each of these circuits. It contains a high voltage transformer, a rectifier, a filter circuit, a small transformer with taps from 1.1 to 32 v and several filament supplies. Mounted on the top of the power pack are 12 tube sockets and each of these is wired for a particular circuit. On the baseboard of each circuit there is a similar tube socket and it can be connected to the corresponding one on the power pack by means of battery cables with prong plugs on both ends. In this way all the needed potentials for any of the circuits can be connected in a few seconds without the danger of wrong connections. These tube sockets vary in the number of prongs from four to seven. In order to have an insulated potential to use as a C-battery for biasing purposes, a 45-v B-battery was connected to the power pack and, therefore, it could be connected to any of the tube bases desired. On the front of the power pack are two twin outlets for 110 v. These are helpful in setting up a circuit because they render extension cords and extra outlets for the oscillograph, electron switch, etc., unnecessary. One end of the power pack has binding posts so that the various potentials can be made available for use in other experiments. This type of power pack has proved very convenient and useful, both in the lecture demonstrations and in the electronics laboratory.

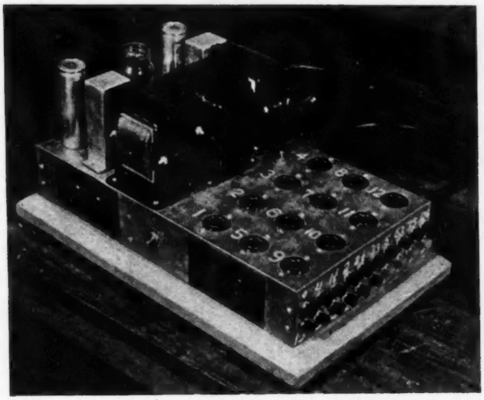


Fig. 11. The power pack.

I wish to acknowledge my indebtedness to my colleague, Mr. T. E. Lothery, Jr., for his help in taking the photographs and to my students, Mr. W. J. Cronin and Mr. W. A. Mussen, for their work in building the apparatus.

Alternating-Current Stroboscope

Grant O. Gale
Department of Physics, Grinnell College, Grinnell, Iowa

AMONG the many interesting phenomena to be studied stroboscopically are a large group of acoustic and other experiments involving frequencies and speeds of the order of several hundred cycles per second. Frequency control and measurement then become an important problem, as they no longer can be carried out mechanically.

Professor Gingrich[1] has described a stroboscope circuit fired by a relaxation oscillator, with a direct-reading impulse counter used up to about 20 pulses/sec, beyond which the oscillator must be calibrated or a direct-reading frequency meter used.[2] While the relaxation oscillator does give a wave form with a steep wave front and is ideally suited for firing the tube, there are times when it is both convenient and desirable to be able to fire the stroboscope circuit with the usual sinusoidal wave form peculiar to alternating-current generators and most audio-oscillators. Accordingly, the following

[1] Gingrich, Am. Phys. Teacher **5**, 277 (1937).

[2] Gingrich, Evans and Edgerton, Rev. Sci. Inst. **7**, 450 (1936).

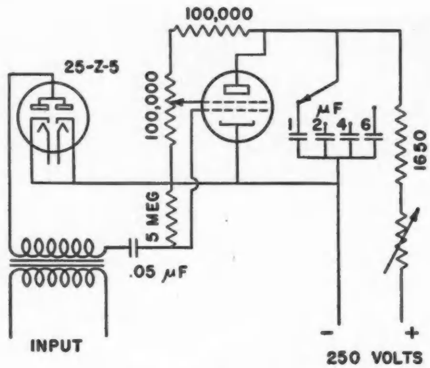


FIG. 1. Diagram of circuit.

stroboscope circuit has been worked out to operate on input signals of sinusoidal wave form, though it has been found that it will also operate satisfactorily with mechanical interrupters and various other input signals.

The tube used is the Strobotron, a special tube with a cold cathode.[3] The duration of the flash is very short, giving distinct stroboscopic patterns; and the de-ionization time of the tube is rather long, so that the upper frequency limit is about 300 flashes/sec.

The circuit shown in Fig. 1 operates satisfactorily on an input signal of about 15 to 20 v, less voltage than this usually causing the tube to fire at a subharmonic. The 25Z5 acts as a half-wave rectifier and thus fires the tube. In order to cover a wide range of frequencies with greater stability several condensers are provided as shown. With 250 v on the plate of the Strobotron, the 6-μf condenser gives best results below 30 vib/sec; the 4-μf condenser, at about 60 vib/sec; the 2-μf condenser, up to 100 vib/sec; and the 1-μf condenser, for the higher frequencies. As the frequency of the discharge increases, the plate current in the Strobotron tube naturally increases, and the variable resistance is provided to keep the plate current below about 30 ma.

Advantages: microphone pick-up.—The chief advantage of the circuit is that, since it is possible to fire it with a sinusoidal wave form, audio-frequencies may be picked up and amplified, and the output of the amplifier used to fire the tube. This possibility has many interesting applications. For demonstration work it provides for exact synchronization—we have so used it with the Melde standing wave experiment. The pick-up from the fork, amplified, fires the stroboscope and either the fork or string may be observed and studied.

Notes from musical instruments may be picked up and the frequency measured by comparison with stroboscopic patterns drawn on calibrated rotating disks.

Frequency measurements.—Many laboratories have variable audio-oscillators[4] which either are direct-reading or can be calibrated by means of the usual frequency bridge. The stroboscope then makes a reliable frequency meter—as reliable as the oscillator firing it. We have used it with a direct-reading Western Electric 13A oscillator—from 20 cycle/sec to above "middle C." This covers a useful range in acoustics and provides a convenient way to measure frequencies of loaded forks, vibrating strings, etc.

Speed measurements.—With a standard oscillator, the stroboscope makes an ideal speed indicator for electric motors, etc. Even high speed motors, such as grinder motors with speeds of about 3600 rev/min, give satisfactory stationary images without the use of multiple patterns. The circuit can also be fired by the 60-cycle power supply (20 v), giving a convenient and fairly stable frequency for studying such effects as the slip of induction motors, the damping of gyroscopes and other rotating parts,

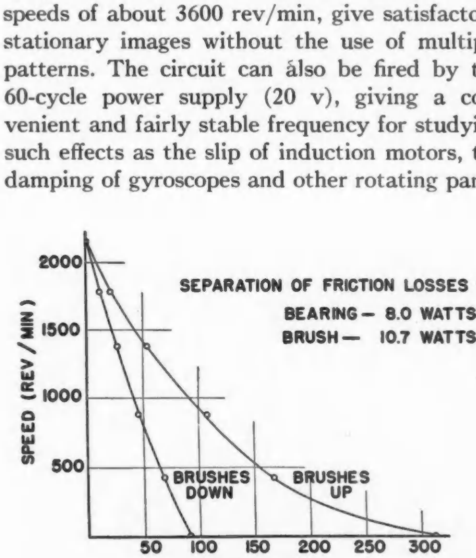


FIG. 2. Retardation curves for a $\frac{1}{3}$-hp motor.

[3] This tube and its characteristics are described by Germeshausen and Edgerton in Electronics, Feb. and March, 1937.

[4] Williamson, Am. Phys. Teacher **5**, 135 (1937).

and the retardation curves of electric motors.[5] Fig. 2 shows the retardation curves for a $\frac{1}{3}$-hp motor with the brushes in place and with the brushes lifted.

The actual construction of the stroboscope was carried out by Mr. Sam Johnston, a senior physics student, in the Department shop. Standard radio parts were used wherever possible. The tube itself was mounted as a separate unit at the center of a parabolic reflector. This unit is light and easily moved about, and is connected to the circuit by a standard 4-conductor cable with plug and socket.

[5] Karapetoff, *Experimental Electrical Engineering* (ed. 1), p. 398.

An Experimental Study of Simple Harmonic Motion

PHILIP A. CONSTANTINIDES
Department of Physical Science, Wright Junior College of the City of Chicago, Chicago, Illinois

AN examination of the most extensively used college physics laboratory manuals and of some representative manuals prepared by institutions for the needs of their own laboratories will reveal the fact that not a single experiment is described or even suggested for the study of the *fundamental* characteristics of simple harmonic motion. I say fundamental characteristics, advisedly, since in all cases where experiments are presented they are limited practically to one aspect of the topic; that is, to verifying experimentally the value of the period of a vibrating body calculated on the assumption that its motion is simple harmonic. This type of experiment does not give insight into the elements of the motion of the body or experimental proof that the motion actually satisfies the dynamical condition of simple harmonic motion.

An elementary but precise method of locating the positions of a vibrating body, such as a simple pendulum,[1] at successive equal time intervals, and thereby proving that its motion is practically simple harmonic is as follows. A heavy ball *B* (Fig. 1) having a metallic stylus attached to the bottom is suspended by means of a long, fine flexible wire about 160 cm in length from a support stand[2] *S*. Under the pendulum, at a distance of about 3 mm from the projecting stylus and concentrically to the arc of vibration of the stylus, is placed a metal strip *M* (Fig. 1) fastened to a wood base. The end *S* of the suspension wire and the metal strip *M* are connected to the terminals of a spark-timer *T* so that when the spark coil circuit is closed, spark traces are produced on a strip of waxed paper lying on the metal strip. With the spark-timer used, the discharge occurred at intervals of 1/30 sec.

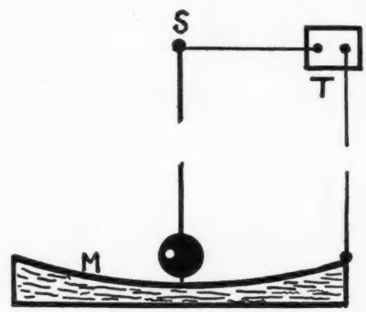


FIG. 1. Diagram of apparatus.

To avoid confusion resulting from possible superimposition of points due to the oscillatory motion of the pendulum, it is advisable to give to the pendulum a slightly elliptical motion, to start the spark discharge when the ball is passing over the position *A* (Fig. 2) and to stop it when the ball, after a complete oscillation, is passing over the position *C*. The traces so obtained have the form of an ellipse of large eccentricity. In one series of determinations, the major axis was about 56 cm and the minor axis 3 cm. The distance

[1] The same method can be used for the study of the motion of a physical pendulum.

[2] The balls and support stand used in the experiment on elastic impact are satisfactory.

between the spark points varied from about 20 mm in the region of equilibrium position to 2 mm in the region of maximum displacement.

At the region of maximum displacement, where the speed is small, two sparks may pass through the same point. Such points are rather easily identified either by their larger diameter or the larger area of melted wax. The position of equilibrium E may be accurately determined by producing a brief discharge when the pendulum is at rest.

Displacement-time curve.—The accurate construction of a curve showing displacement as a function of time is the first objective of the experiment since this enables the student to ascertain how the velocity and the acceleration of the pendulum vary during the oscillation. The time corresponding to a given point P (Fig. 2) is determined by counting the number of spark intervals from this point to a point A near the minor axis OO' of the ellipse taken as the origin

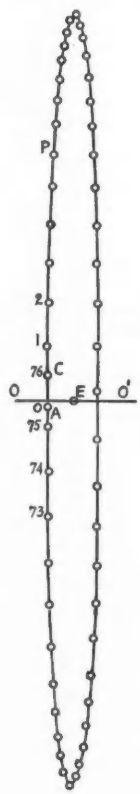


FIG. 2. Spark trace of pendulum's motion.

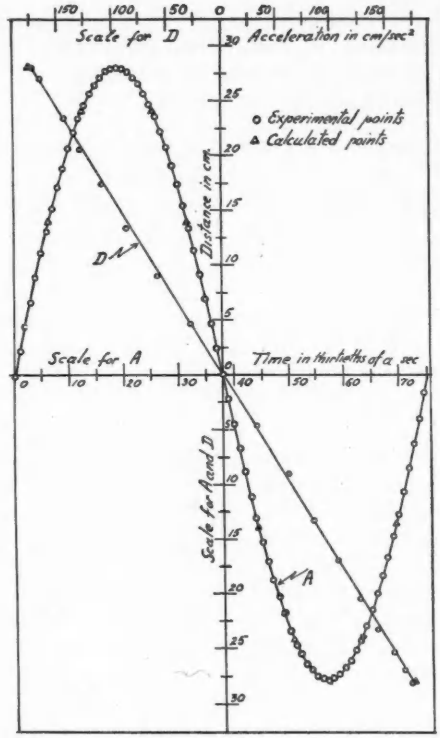


FIG. 3. A, displacement-time curve; D, acceleration-displacement curve.

of time. The displacement of P is its distance from the axis OO'. The curve A, Fig. 3, based on the data appearing in Table I, fits beautifully and convincingly with the theoretical curve obtained from the relation

$$s = A \sin \theta = 28 \sin \theta,$$

where A ($=28$ cm) is the amplitude of the pendulum. The triangular points along the curve A indicate the theoretical values of the curve for every 30° beginning with zero displacement.

The velocity-time curve.—From the displacement-time curve other curves, such as the *velocity-displacement* or the *velocity-time* curve, can be obtained. The velocity at the instant corresponding to the point P, Fig. 4, is found from the slope of the displacement-time curve at the point P; or

$$v_P = ds/dt = L_1M_1/M_0M_1.$$

TABLE I. *Observed times t and displacements s.*

t (1/30 SEC)	s (CM)	t (1/30 SEC)	s (CM)	t (1/30 SEC)	s (CM)
0	0.30	26	23.55	51	24.50
1	2.00	27	22.18	52	25.45
2	4.35	28	20.70	53	26.42
3	6.58	29	19.08	54	26.98
4	8.80	30	17.35	55	27.46
5	11.05	31	15.43	56	27.70
6	13.17	32	13.38	57	27.85
7	15.18	33	11.32	58	27.63
8	17.05	34	9.13	59	27.35
9	18.85	35	6.88	60	26.88
10	20.55	36	4.62	61	26.13
11	22.05	37	2.35	62	25.23
12	23.30	38	0.00	63	24.25
13	24.50	39	2.25	64	23.05
14	25.57	40	4.62	65	21.60
15	26.43	41	6.73	66	20.13
16	27.15	42	8.91	67	18.45
17	27.66	43	11.20	68	16.67
18	27.88	44	13.20	69	14.73
19	28.00	45	15.30	70	12.88
20	27.95	46	17.08	71	10.63
21	27.57	47	18.80	72	8.45
22	27.17	48	20.40	73	6.33
23	26.55	49	21.90	74	4.10
24	25.73	50	23.28	75	1.78
25	24.63				

The maximum velocity v_m is an important value in the plotting of the *velocity-time* curve. Since v_m occurs when the pendulum passes through the position of equilibrium, its value can be obtained with precision, for the direction of the tangent at the point of inflection of the *displacement-time* curve can be determined accurately. From Fig. 4 we have

$$v_m = (L_2M_2 + L_3M_3)/M_2M_3.$$

From curve A, Fig. 3, we obtain $v_m = 2.32$ cm/(1/30 sec) = 69.6 cm sec^{-1}. This value can be compared with the value computed from the equation

$$v_m = A(K/m)^{\frac{1}{2}},$$

where K is the force constant and m is the mass of the vibrating body. In the case of the simple pendulum

$$v_m = A(g/l)^{\frac{1}{2}} = 2\pi A/T;$$

and, for $T = 2.520$ sec, v_m is 69.9 cm, which differs from the observed value by 0.43 percent.

It is true that in all cases such precision is not obtainable since there is always some element of doubt about the precise direction of the tangent; but with graphs of large scale and with some training it will be found that the graphically determined values for maximum velocities do not differ much more than 0.5 percent from the theoretical values. The graphically determined *velocity-time* curve B, Fig. 5, fits satisfactorily with the theoretical curve given by

$$v = v_m \cos\theta = 69.9 \cos\theta.$$

Acceleration-time curve.—The value of the acceleration at the instant corresponding to any point is found from the slope of the *velocity-time* curve at the point P. The values so obtained must, of course, be multiplied by 30 in order to obtain the acceleration in centimeters per second per second. In this case the experimentally determined maximum acceleration a_m was found to be 171 cm sec^{-2}. This value can be compared with the theoretical value given by the relation

$$a_m = AK/m = Ag/l = 4\pi^2 A/T^2.$$

From this experiment

$$a_m = 4 \times (3.14)^2 \times 28.0/(2.52)^2 = 174 \text{ cm sec}^{-2}.$$

The difference between the two values is 1.7 percent and this represents the average degree of precision attainable in the experimental determination of the acceleration. The graphically determined curve C, Fig. 5, fits satisfactorily with the theoretical curve given by $a = a_m \sin\theta$.

Displacement-acceleration curve.—When the displacements are plotted as a function of the accelerations (curve D, Fig. 3), it is seen that the points lie approximately on a straight line

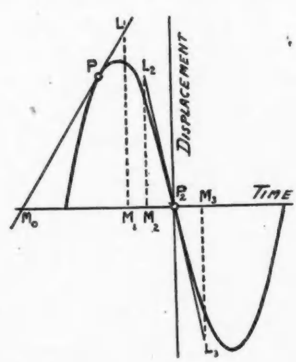


FIG. 4. Method for finding velocities.

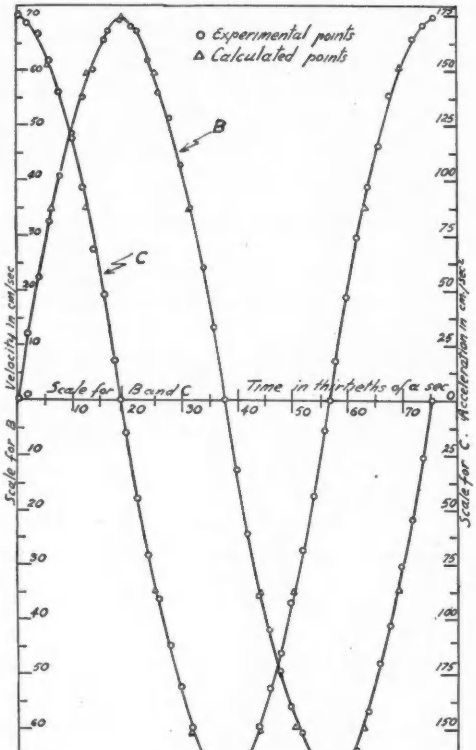


FIG. 5. *B*, velocity-time curve; *C*, acceleration-time curve.

passing through the origin. Thus, the experiment shows that the characteristic feature of the motion of a simple pendulum is that its acceleration is proportional to the displacement, or what is the same thing, that the force producing the motion is proportional to the displacement and directed towards the position of equilibrium. We conclude, therefore, that the motion of the pendulum of small amplitude is simple harmonic, since it satisfies the characteristic dynamical conditions of such motion.

Supplementary experiment.—From the trace measurement it was found that the period was equal to $75+(14.3/23.5)$ time intervals, or 2.520 sec. The value of the period obtained from measurements of the length of the pendulum (156.70 cm) on the assumption that the amplitude of the vibration was very small is[3]

$$T_0=2\pi(l/g)^{\frac{1}{2}}=2\pi(156.70/980.27)^{\frac{1}{2}}=2.51 \text{ sec.}$$

When corrected for vibrations of an angular amplitude of 10°, the period is 2.52 sec.

This value is the same to three significant figures as the value determined from the tracing, which implies a variation of less than 0.2 percent from the experimentally determined value. On the other hand, the value of T determined with a stopwatch on the basis of the time measurements for 60 vib was found to be identical with the value obtained by means of the traces. The foregoing values are typical of the results obtained with other traces and give a definite idea of the degree of precision of time measurements attainable by this method. More specifically, the greatest variation of the period from the average value of 10 readings of approximately the same amplitude was found to be less than 0.003 sec. or about 0.12 percent of the period. This result is consistent with the magnitude of error calculated on the assumption that the variation of the distance between successive points due to the uncertainty of the path of the sparks, may amount to as much as 2 mm, that is, to about 1/10 of the total distance between two sparks in the neighborhood of the position of equilibrium of the pendulum. In that region, with the amplitudes considered, a distance of 1 mm corresponds to about 0.0015 sec.

In case the complete determination of *velocity-time* and *acceleration-time* curves appears long and tedious for some groups of students, I enumerate in Table II some of the determinations that the student may make from the study of the *displacement-time* curve.

TABLE II. *Features of the displacement-time curve.*

ASPECT OF CURVE	SIGNIFICANCE
Abscissa	Time (determination of period)
Ordinate	Displacement (type of motion)
Slope	Velocity
Steepest slope	Maximum velocity (easily determined)
Horizontal slope	Minimum velocity
Sharpest curvature	Maximum acceleration

A similar method has been tried in connection with the study of the motion of a weight suspended from a spring. However, the interpretation of the results is more complicated because the body vibrates on practically the same straight line so that the determination of the succession of the points at the extremities is not easy.

[3] The value used for g, namely, 980.27 cm sec^{-2}, has been corrected for the altitude and latitude of the locality where the experiment was performed.

A Miniature Kundt Tube

Howard S. Seifert
Department of Physics, Kalamazoo College, Kalamazoo, Michigan

A STRIKING demonstration of stationary sound waves in the range just above the audible region may be made with the help of a Galton whistle and a small amount of lycopodium powder. About 20 to 50 mm^3 of lycopodium is spread out in a line (by tilting and tapping) inside a glass tube of about 7-mm inside diameter and any convenient length between 10 and 60 cm. The Galton whistle (in our case, a Cenco No. 85155 whistle) is placed close to the open end of the tube with its orifice so oriented that the air jet is directed perpendicular to the tube rather than into it. The opposite end of the tube is closed with a smooth cork stopper and the whistle adjusted to the upper part of its frequency range.

If the whistle is blown with an air pressure of about 6 cm of mercury, to provide ample energy, it is possible upon slowly tuning the whistle to observe turbulent ridges in the lycopodium similar to those produced in the usual large-scale Kundt tube. With careful tuning of the whistle, and with a slight rotation and tapping of the glass tube, these antinode ridges form thin semicircular, paper-like partitions[1] only a few hundredths of a millimeter thick, which are quite stationary and may be located with considerable accuracy. Thus it is possible in a tube about 50 cm long to measure very short wave-lengths with an accuracy greater than that obtainable with the wave-lengths corresponding to the audible region which are produced in the usual undergraduate experiment with the Kundt tube.

The experiment is more impressive if the whistle resonance chamber is made 1 mm or less in length, when it will emit inaudible or scarcely audible frequencies of 16,000 vib sec^{-1} or higher, and the distance between loops will be approximately 1 cm. In Fig. 1 the distance between loops was 0.93 cm and the frequency (using $\nu = v/\lambda$) was approximately 18,600 vib sec^{-1}. The fine ripples between the principal ridges are explained by Andrade[1] as due to vortex motion of the air in the tube. The delicacy of the pattern was limited only by the Galton whistle, which, because of its design (Fig. 2), permitted the resonating column a minimum volume of about 8 mm^3. The smaller patterns are more clean-cut if produced in short tubes, since the longer air columns possess many more closely adjacent high-frequency resonance points than the shorter ones.

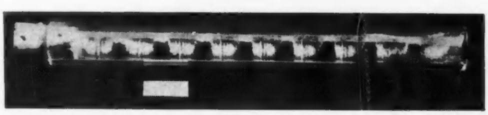


Fig. 1. Lycopodium loops produced by stationary waves of frequency 18,600 vib sec^{-1}. The white rectangle is 1 cm long, for comparison. When sound waves act on the tube, the central ridges are much sharper than shown here.

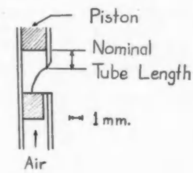


Fig. 2. Showing how the minimum volume of the resonance cavity is limited by the design of the Galton whistle.

An approximate check on the whistle frequency was obtained (Fig. 3) by picking up the sound of the whistle on a crystal microphone and feeding the amplified output into an oscilloscope with a calibrated sweep circuit. The calibration of the oscilloscope was checked with the aid of the whistle itself. The whistle was tuned to a high multiple (say 10 or 20 times) of the frequency of a standard tuning fork of frequency 512 vib sec^{-1} by observation of the oscilloscope patterns. The whistle pressure was then carefully held constant while the oscilloscope pattern was set at various simple ratios, giving calibration points on the sweep circuit dials. It is easier to calibrate an oscilloscope by using frequency standards *higher* than the sweep frequency than it is to use lower frequencies. Some typical frequencies and pattern sizes obtained in a tube 8.8 cm long are listed in Table I.

It was impossible to check frequencies even

[1] See Andrade, Phil. Trans. **A230**, 413 (1932), especially p. 418.

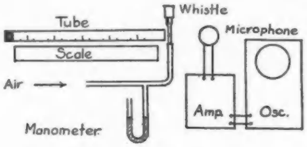


FIG. 3. General arrangement of apparatus when checking resonance frequencies with oscilloscope.

approximately by applying the simple closed tube formula to the resonating chamber of the Galton whistle. The empirical wave-length correction term K in the simple equation $\lambda = 4l + K$ varies inversely with both the pressure and the whistle length l. At the higher frequencies K was larger than $4l$. For example, with $l = 0.236$ cm and $p = 2.3$ cm of mercury, K was 1.8 cm, almost twice the value of $4l$. Here l is taken as the distance from the movable piston to the upper edge of the whistle opening. If l is measured, instead, to the lower edge, an increase of 2 mm, the values of K are somewhat reduced but still vary in the same way and are much too large to be neglected. To obtain steady reproducible frequencies with a Galton whistle the pressure

TABLE I. *Typical characteristics of miniature Kundt tube with Galton whistle*

Net Pressure (cm-of-Hg)	Nominal Whistle Length (cm)	Actual Measured Wave-Length (cm)	Frequency (vib sec^{-1}) Pattern	Oscilloscope
3.5	0.962	5.00	6,900	6,600
2.4	0.609	4.00	8,620	8,000
3.5	0.378	3.11	11,100	10,400
2.3	0.236*	2.69	12,800	12,000
4.4	0.251*	2.65	13,030	12,800
4.4	0.163	2.33	14,800	14,700
4.4	0.092	2.09	16,500	16,600
6.2	0.049	1.85	18,620	18,650

must be carefully regulated, as an increase in pressure of 100 percent can cause an increase in frequency of 20 percent. It will be noticed in Table I that, in the case of the two starred values, the frequency actually *increased* as the whistle was made *longer*, due to a doubling of the pressure.

While the demonstration of inaudible sound waves in this manner is not an exact experiment, it provides a simple and effective qualitative demonstration.

Appointment Service

REPRESENTATIVES of departments or of institutions having vacancies are urged to write to the Editor, Columbia University, for additional information concerning the physicists whose announcements appear here or in previous issues. *The existence of a vacancy will not be divulged to anyone without the permission of the institution concerned.*

Positions Wanted

26. Ph.D., with long experience in an American college in China, wishes a college teaching position.

27. Ph.D., physics, Northwestern '35; A.B., engineering, Harvard. Age 42, married, 3 children. Experience: 1 yr, lt., artillery; 12 yrs business and sales; 5 yrs college teaching. Interested in undergraduate teaching, including astronomy.

29. Ph.D., Northwestern; M.S., Pittsburgh; A.B., Muskingum. Age 34, married, 1 child. Has had 13 yrs teaching experience in two universities. Interested in teaching and research.

30. Ph.D., Univ. of Chicago. Many years experience as head of department of physics in prominent college. Author of books on physics and history of science. Large work on history of physics in preparation. Interested in college or university teaching.

31. Ph.D., Columbia. Years of experience as head of departments of physics in colleges and universities. Author of new type of laboratory manual. Designer of many new types of simplified apparatus. Research in radio, acoustics and methods of teaching physics.

33. M.S., experimental physics, coupled with thorough background of courses in professional education. Has taught physics and mathematics for 3 yrs in large high school. Desires position as instructor in high school physics in a university or college experimental or training school.

34. Ph.D., M.S., Penn State. Age 38, married. 13 yrs teaching experience in colleges and universities; 3 yrs head of department in small college; industrial research experience. Interested in teaching, research and administrative work in a small college.

Departments having vacancies or industrial concerns needing the services of a physicist are invited to publish announcements of their wants; there is no charge for this service.

Any member of the American Association of Physics Teachers may register for this Appointment Service and have a "Position Wanted" announcement published without charge.

NOTES AND DISCUSSION

A Simple Device for Demonstrating the Components of a Vector

A SIMPLE device has been constructed for demonstrating the components of a vector.[1] Built as a single unit, it saves much time in setting up the class demonstration. A thin wooden arrow, 40×1.5 cm, is mounted (Fig. 1) on a board 75×85 cm and made to rotate about one end by means of a double lever arrangement on the back of the mounting (Fig. 2). To one end of this double lever is attached a second, slotted arrow, also 40×1.5 cm, which is arranged to slide horizontally (Fig. 1). The tip of a third arrow is attached to the tip of the rotating arrow and is kept vertical by means of a wire mounted on the back of it which passes through a ring in the end of the horizontal arrow. The arrows are painted white to contrast with the mounting. As the horizontal arrow moves to the left and the vertical arrow rises, that part of the horizontal arrow to the left of the axis of rotation disappears under a suitably mounted board. By a similar arrangement, that part of the vertical arrow below the horizontal arrow is kept out of sight. The slotted arrow could be arranged to slide in any desired direction, other than horizontal, to demonstrate components of a vector that are not at right angles.

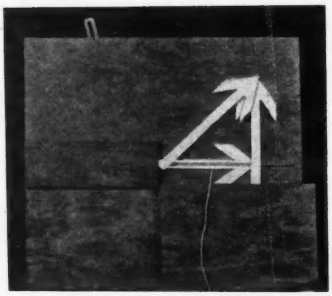


FIG. 1. Photograph of the device.

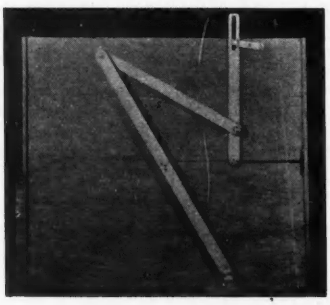


FIG. 2. Rear view of the device.

When the first arrow is rotated the class can watch one component grow as the other decreases. When one component is equal to the magnitude of the vector and parallel to it, the other is zero.

I wish to express my thanks to Mr. Howard I. Pratt, who constructed the device.

WILFRID J. JACKSON

N. J. C. Rutgers University,
New Brunswick, New Jersey.

[1] Sutton, *Demonstration Experiments in Physics* (McGraw-Hill, 1938), p. 16.

A Variation of Kundt's Method for the Speed of Sound

IN *A Textbook of Sound*,[1] A. B. Wood describes a phenomenon in the Kundt tube which can easily be used for measuring the speed of sound, but which seems to be little known in this country. This is the antinodal ring. The phenomenon seems to have been first reported by Dvorak,[2] and later by Andrade and Lewer.[3,4] The last two authors employed essentially the arrangement described below and they suggested the use of the antinodal rings for measuring the wave-length and speed of sound. The apparatus is found in practically all laboratories today. The rings are not difficult to produce, are not over a millimeter in thickness and are stable in position if the tube fittings are tight.

Considerable energy is needed. The writer uses an oscillator of frequency about 1000 cycle/sec., the output being filtered, put through two stages of amplification and fed into a loudspeaker operated at almost full volume. A plate with a nozzle about 3 cm in diameter is fitted over the face of the loudspeaker. The tube in which the rings are produced is of Pyrex, 32 mm outside diameter and about 110 cm long. Over one end is cemented a short length of brass tube which fits tightly over the output nozzle of the loudspeaker. Since glass tubes vary somewhat in diameter, another brass tube of the same diameter and about 25 cm long is cemented to the other end of the glass tube and fitted with a plane-faced piston which fits snugly in the tube. Cork dust which will pass an 80-mesh sieve is placed all along the bottom of the tube. Coarser dust has been found not to work well. By proper adjustment of the length of the air column, sharp striations of the dust may be formed, apparently one particle thick and of varying heights. The antinodal rings form best when the adjustment is such as to produce turbulence in the tube. The frequency of the exciting current can be found by various bridge methods, for example, the series resonance bridge[5] or the Campbell frequency bridge.[6] With a frequency slightly above 1000 cycle/sec., six rings can be formed in a tube 110 cm long. Careful work on the part of students has

usually resulted in values of the speed of sound within 1 percent of the accepted value for the existing temperature.

HARRY E. HAMMOND

University of Missouri,
Columbia, Missouri.

[1] Wood, *Textbook of Sound* (Macmillan, 1930), p. 180.
[2] Ann. d. Physik **153**, 102 (1874); **157**, 42 (1876).
[3] Nature **124**, 724 (1929).
[4] J. Sci. Inst. **7**, 52 (1930).
[5] B. Hague, *A. C. Bridge Methods* (Pitman, ed. 2, 1930), p. 270.
[6] Reference 5, p. 325.

A Modified Cryophorus

THE cryophorus illustrated in Fig. 1 was developed to show clearly, in as short a time as possible, the freezing

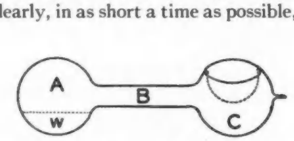


FIG. 1. A modified cryophorus.

of water by evaporation and, in addition, to show that the freezing is not produced by loss of heat by conduction to the refrigerant. Incidentally, the design lends itself to vertical projection of the freezing process. This projection becomes especially striking and beautiful when crossed Polaroid disks are used.

The over-all length of the cryophorus is about 25 cm. The quantity of distilled water W used for the filling is about 15 ml. When the depression in C is filled with a slush of dry ice and alcohol or ether, the water in the other bulb A freezes in less than 30 sec. Our record time is 9 sec.

The freezing time is not materially affected if the connecting tube B between the bulbs is heated with boiling water or if it is gently flamed.

The ice may be quickly melted, for repetition of the experiment, by dumping the refrigerant and pouring hot water over the refrigerating bulb C.

The author wishes to thank the Central Scientific Company for assistance in the construction of the device.

ROSS A. BAKER

Department of Chemistry,
College of the City of New York,
New York, N. Y.

An Explicit Name for the Electrostatic Unit of Charge

MOST elementary textbooks begin the study of electricity with the electrostatic system of units. The unit of charge is abbreviated to "esu," which is short to write and easy to pronounce. Unfortunately, when the concepts of field intensity and potential are introduced, both are given the same name; this is noticeable especially in the statement of problems. Many students conclude that all three quantities have the same unit; this adds to the difficulty of grasping the nature of these new concepts, so unlike anything previously studied in the physics course.

The term *e-s coulomb*, used in intermediate courses, is not available because the term *coulomb* has not been introduced; besides, it is a long expression to use orally.

I have found it useful to introduce a definite name for the electrostatic unit of charge, even though it is not expected to become part of the student's semipermanent vocabulary of physics. The name used is the pronounceable *esuch*, of obvious derivation. The advantage lies in the fact that field intensity is expressed as *dynes per esuch*, and potential as *ergs per esuch*. The practice of canceling units can be continued; for example, the product of potential difference and charge is shown to be *ergs per esuch times esuch*, or ergs. The student cannot miss the fact that this product represents work or energy.

ROBERT S. SHAW

College of the City of New York,
New York, N. Y.

Reprints of Survey Articles for Class Use

Reprints of the following survey articles which have appeared in various issues of *The American Physics Teacher* may be obtained from the Editor, Pupin Physics Laboratories, Columbia University, New York, New York. Stamps will be accepted in payment.

L. A. DuBridge, *Some Aspects of the Electron Theory of Solids*. 50 cts. for 6 copies.
R. T. Birge, *The Propagation of Errors*. 50 cts. for 6 copies.
Merle Randall, *Electrolytic Cells*. 60 cts. for 6 copies.
J. C. Stearns and D. K. Froman, *Cosmic Rays—Their History, Source, Nature and Effects*. 60 cts. for 6 copies.
W. G. Cady, *A Survey of Piezoelectricity*. 60 cts. for 6 copies.
C. W. Ufford, *Spectroscopy—A Survey*. 60 cts. for 6 copies.
Dean E. Wooldridge, *The Separation of Isotopes*. 60 cts. for 6 copies.
Committee of the American Physical Society, *Physics in Relation to Medicine* (Reprint of 1923 report). 10 cts. per copy.

RECENT PUBLICATIONS

First-Year Textbooks

Household Physics. Walter G. Whitman, Department of Physical Science, State Teachers College, Salem, Massachusetts. 443 p., 1 plate, 30 tables, 412 fig., 15×23 cm. *Wiley*, $3. The present edition of this well-known textbook retains the general plan of the 1924 and 1932 editions. Chapters that have been expanded or brought up-to-date include those on electrical devices, illumination and radio. Six new chapters have been added, their titles being Protection against fire, Air-conditioning, Electricity in everyday life, Visual aids, The camera and photography, and Solutions and other dispersions. A wealth of modern applications of physics in the home and in everyday life are described in this book. The photographs and diagrams are appropriate and informative.

Physics for Science Students. Harold C. Barker, Professor of Physics, University of Pennsylvania, and Melvin R. Harkins, Professor of Physics, University of Pennsylvania. 516 p., 413 fig., 14×22 cm. *Pitman*, $3.50. This text is intended for a second-year general course for engineering and science students who have at least begun the study of elementary calculus. The authors have deliberately made the treatment very concise, and the selection and treatment of the material conventional. For brevity, many of the illustrative remarks ordinarily found in general texts are omitted. In the order of subjects, light follows wave motion and sound, and precedes heat. A list of problems with answers accompanies each of the five main sections. No reading or literature references are given.

Advanced Physics

Electron Optics. Compiled by Otto Klemperer, Electric and Musical Industries (England). 117 p., 48 fig., 14×22 cm. *Cambridge Univ. Press and Macmillan*, $1.75, paper cover. The fourth in the series of *Cambridge Physical Tracts* [Am. Phys. Teacher **7**, 139 (1939)], this monograph gives a concise introductory account of important principles, methods and applications of geometrical electron optics, based on both the general literature of the field and reports of researches performed in the E. M. I. laboratories. It is intended for advanced students and research workers, but a special knowledge of geometrical optics or electron physics is not presupposed.

Electron Optics—Theoretical and Practical. L. M. Myers, Research Department, Marconi's Wireless Telegraphic Co. Ltd. 636 p., 379 fig., 1 plate, 14×23 cm. *Van Nostrand*, $12. A comprehensive and detailed treatment of the theoretical, experimental and practical aspects of electron optics is here provided that should be of great value to the graduate student who intends to specialize in this branch of vacuum physics. It provides a treatise in English that is similar in general character and importance to the excellent German book, *Geometrische Elektronenoptics* (1934), by Brüche and Scherzer, but is, of course, more up-to-date, with its accounts of the great amount of work carried on during the last few years by such active groups as Brüche and his collaborators in Germany, and Zworykin and his co-workers in the United States.

History and Biography

Portraits of Eminent Mathematicians, with Brief Biographical Sketches. Portfolio No. 2. David Eugene Smith, Professor Emeritus of Mathematics, Columbia University. *Scripta Mathematica* (Yeshiva College, New York, N. Y.), $3. Similar in character and quality to Portfolio No. 1 [Am. Phys. Teacher **4**, 216 (1936)], the present volume contains reproductions of portraits of Euclid, Cardan, Kepler, Fermat, Pascal, Euler, Laplace, Cauchy, Jacobi, Hamilton, Cayley, Chebishef and Poincaré. Each portrait is 25×35 cm, and is loosely slipped into a folder containing a biographical sketch written by Professor Smith. It is of interest that copies of these portfolios were selected by a committee of scientists for inclusion in the Time Capsule buried in the grounds of the New York World's Fair. Portfolio No. 3, for which advanced subscriptions are now being received, will be devoted to physics; it will contain portraits of Newton, Galileo, Huygens, Ampère, Fresnel, Faraday, Joule, Clausius, Maxwell, Gibbs, Hertz and Rowland, with biographical sketches prepared by Professor Henry Crew.

Miscellaneous

Supersonics, the Science of Inaudible Sounds. Robert Williams Wood, Professor of Experimental Physics, Johns Hopkins University. 168 p., 42 fig., 12×19 cm. *Brown University Press*, $2. The three interesting and informative popular lectures contained in this small volume deal with high frequency sound waves, their mechanical and electrical sources, and their physical and biologic effects. The phenomena and experimental aspects of the subject are stressed.

Your Automobile and You. Roy A. Welday, Instructor in Physics, Scott High School, Toledo, Ohio. 264 p., 108 fig., 12×19 cm. *Henry Holt*, 88 cts. Designed for use as the basis of a practical secondary school course in automobile driving, this intelligently planned text contains very simple but complete instructions for those who have never driven a car before, and much information that should be of interest and value to the average driver of motor vehicles. The main topics discussed are: brief history of vehicles; parts, physics and care of automobiles; physical fitness and psychology of drivers; highway rules, customs and laws; accidents and insurance. The book is intended for a half-year course, but the author provides study guides which show how it can also be used for ten-week, or even two-week, short courses.

DIGEST OF PERIODICAL LITERATURE

Apparatus and Demonstrations

Some useful demonstration apparatus. N. H. Black; *Sch. Sci. and Math.* **38**, 705–707, June, 1938. Among the 10 pieces briefly described are the following: (1) a working model of a 3-element vacuum tube, magnified 6 times, the grid being a spiral wire and the plate which surrounds it being cut away to render visible the glowing filament within; (2) a sodium flame, intense enough to make possible the projection of absorption lines of sodium before a large class, produced with a bit of sodium placed in a small stainless steel boat over a Bunsen flame and inside a cylindrical chimney; (3) an "optical disk," suitable for demonstrations before a large class, consisting of an ordinary optical bench, an arc lamp, a large condenser, a screen with several parallel slits, and various 6-in. blocks of glass to serve as refracting pieces; (4) a device for keeping the soap film from breaking during a demonstration of interference bands produced by soap films; a small crystallizing dish containing the soap solution is placed in a 5-in. cubical box which is equipped with a glass window slightly tilted to prevent reflection from it; a ring is mounted on an axle tangential to its periphery so that it can be dipped into the soap solution and then brought into a vertical plane; the atmosphere inside the box soon becomes saturated, which lengthens the life of the film.—D. R.

Transparent projections of lecture experiments. W. J. Conway; *J. Chem. Ed.* **16**, 314–316, July, 1939. The activity series of the metals, supersaturation and other lecture experiments may be demonstrated by using an ordinary lantern slide projector and glass cells thin enough to produce good definition of their contents. A wooden frame to hold the cells is substituted for the ordinary slide holder. The cells are made from standard lantern slides, $8.3\times10.2\times0.125$ cm. Three strips of glass, about 1 cm wide and 1 to 5 mm thick, cut from thin lantern slide cover glass or from heavier glass, are placed on a slide as shown in Fig. 1 and covered with a second slide. The binding material is polyvinyl acetate (Vinylite, grade AYAA, Carbide and Chemical Corp., New York City) dissolved in acetone to make a 25-percent solution and applied to the slides so as to make contact with the separatory strips. After the acetone has evaporated at room temperature, the cell is assembled and placed under a small weight in an oven at 100°C for 10 min. If the cell leaks, dry it and paint the edges with the solution of polyvinyl acetate. A second cell containing a 1-percent solution of copper sulfate which has been boiled to remove dissolved air may be placed between the first cell and the source of light to absorb the heat.

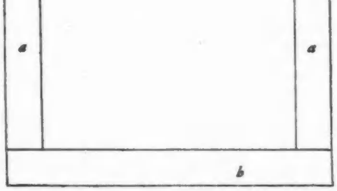


Fig. 1. Diagram of glass cell.

To demonstrate innoculation of a supersaturated salt, heat a 40-percent solution of C.P. sodium acetate to about 85°C until all the salt dissolves. Transfer the warm solution by pipet to a cell heated by partial immersion in warm water. After the cell cools, place it in the projector, focus the meniscus and innoculate. Long, beautifully shaped needle crystals immediately grow completely across the screen. By removing the second, heat-absorbing cell the crystals may be redissolved.

If several cells containing acid are placed successively in the projector and a representative metal is added to each, the relative evolution of hydrogen gives a rough estimate of the activity of metals in the electrochemical series. Divided cells, made by placing several strips of glass parallel to *a*, Fig. 1, allow the demonstration of the activities of several metals at the same time.

For the projection of a lead tree a cell with separating strips 1.25 mm thick is used. The bottom strip *b* is made of lead and has a protruding end; it serves as anode and is made either by folding several thicknesses of lead foil and squeezing in a vice or by hammering down $\frac{1}{8}$-in. sheet lead. A piece of platinum foil inserted at the top of the cell serves as cathode. To a 5-percent filtered agar solution add lead acetate to make a 10-percent solution. Use a funnel, made by drawing out a test tube, to fill the cell. In the thin cell the crystals of the lead tree are in an approximate plane and the entire tree is in focus during its formation.—H. N. O.

Standard of Concert Pitch

International standard of concert pitch. G. W. C. Kaye; *Nature* **143**, 905–906, May 27, 1939. The question of the international standardization of concert pitch was the subject of an international conference held under the auspices of the International Standards Association in London, in May. France, Germany, Great Britain, Holland and Italy sent delegates, and the official views of Switzerland and the United States were before the Conference. It was unanimously agreed to forward to the parent International Acoustics Committee and other committees concerned the following recommendations:

(1) That the international standard of concert pitch shall be based on a frequency of 440 cycle/sec for the note A in the treble clef.

(2) That this value shall be maintained within the closest limits possible by soloists, orchestras, choirs, etc., throughout all musical performances, and also in recorded music.

(3) That, with a view to reducing the necessary tolerances to acceptable values, a set of technical recommendations be drawn up preferably on the basis of international cooperation.

Questions for future study were suggested, such as: the influence of temperature on orchestral instruments and consequent tuning problems; the feasibility of transmitting the standard pitch by broadcasting or telephone; the provision of sub-standards; the production of specifications dealing with tolerances in manufacturing of musical instruments; the setting-up of supervisory safeguards, national and international, for verifying observance of the standard pitch to prevent its falling into desuetude as did the Vienna standard of 1885.—H. N. O.

Combination Tones

Combination tones in sound and light. W. Bragg; *Nature* **143**, 542–545, Apr. 1, 1939. Combination tones in sound, first described by Sorge in 1745 and Tartini in 1754, were the subject of a vigorous disputation between Helmholtz, who claimed they caused a response in tuned resonators, and Koenig, who claimed they originated in the ear. It is now known that they must under certain conditions have an objective existence but, as Helmholtz agreed, can be generated within the ear. Analogous effects are observed in a modulated radio wave and in the Raman effect.

When two vibrations of frequencies ν_1 and ν_2 are imposed on a medium, combination tones of frequencies $\nu_1+\nu_2$ and $\nu_1-\nu_2$ appear only if the amplitude of one source depends on the amplitude of the other; that is, only if a term of the form $\sin 2\pi\nu_1 t \cdot \sin 2\pi\nu_2 t$ appears in the expression for the disturbance of the medium. To demonstrate this, two flat coils, one within the other, through which alternating or direct current may be sent, are placed with their planes at right angles and their centers coinciding. The second coil is free to turn about the line which is the intersection of their planes, and will vibrate if one or both of the currents is alternating. To detect this vibration, a disk is placed on the end of a rod extending out from the coil and rigidly fastened to it. This disk vibrates above a cylindrical tube with a slider containing a piezoelectric microphone connected to a loudspeaker. The tube is tuned by moving the slider. For alternating currents of frequencies ν_1 and ν_2 in the two coils, the torque on the second coil is proportional to $\sin 2\pi\nu_1 t \cdot \sin 2\pi\nu_2 t$, and resonance is observed at lengths corresponding to the sum and difference frequencies, with the difference tone stronger, as theory predicts. Energy transformations between source and speaker alter the relative intensities and introduce new combination tones, but the verification of the theory is clear.

The effect of the multiplication of harmonic terms may be shown in a second way. Light passing through a small square aperture in a screen is focused at a place where it is partly intercepted by two tuning forks at right angles to each other and parallel to the sides of the square. When the tuning forks vibrate with frequencies ν_1 and ν_2, the light is restricted to a rectangle of sides $a+b\sin 2\pi\nu_1 t$ and $c+d\sin 2\pi\nu_2 t$; hence the expression for the light intensity contains the term $bd\sin 2\pi\nu_1 t \cdot \sin 2\pi\nu_2 t$. The light falls on a photoelectric cell, is transformed into sound and analyzed as before; frequencies ν_1, ν_2, $\nu_1+\nu_2$ and $\nu_1-\nu_2$ are found. When the tuning forks are parallel the intensity is proportional to $a+b+c\sin 2\pi\nu_1 t+d\sin 2\pi\nu_2 t$ and no combination tones are observed except faint ones due to the detecting apparatus.

If a piezoelectric cell were inserted at the end of a short cylinder equipped with a piston executing simple harmonic motion, the expression for the volume of the air might be written, $V=V_0+V'\sin 2\pi\nu t$. Under isothermal conditions (the general result is the same for adiabatic) PV is constant and the expression for the pressure P, when expanded, contains harmonics. With two pistons a system of combination tones appear in the expansion. Thus are combination tones generated by detecting apparatus. When a tuning fork is held above the resonating tube in the first experiment, the octave overtone is found. Combination tones of a harmonium at the Royal Institution are greatly intensified when a microphone connected to a loudspeaker is placed on the instrument; the pistons of the aforementioned model are replaced by the surges of the air from pipes or reeds and the combination tones appear to be of greater importance in the wood vibrations than in the air outside.—H. N. O.

Check List of Periodical Literature

Contributions of mathematics to research in physics and chemistry. S. Dushman; *J. Eng. Ed.* **39**, 716–24, May, 1939.

Physics of flames and explosions of gases. B. Lewis and G. v. Elbe; *J. App. Phys.* **10**, 344–59, June, 1939.

Recreating geological history with models. M. B. Dobrin; *J. App. Phys.* **10**, 360–71, June, 1939. Examples of the application of dimensional theory to geologic models.

Our knowledge of atomic nuclei. G. P. Harnwell; *J. Frank. Inst.* **227**, 443–59, Apr., 1939.

What has become of reality in modern physics? W. F. G. Swann; *J. Frank. Inst.* **227**, 473–96, Apr., 1939.

The Doppler effect in astronomy. A. Harvey; *Sch. Sci. Rev.* **20**, 530–44, June, 1939. A summary of applications of the Doppler effect in astrophysics.

Science and the metallurgical industry. J. Johnston; *Sci. Mo.* **48**, 493–503, June, 1939. A general account, with emphasis on the contributions of Willard Gibbs, by the Director of Research, United States Steel Corporation.

Eminent men. M. Smith; *Sci. Mo.* **48**, 554–62, June, 1939. A concise summary of existing knowledge concerning the characteristics of eminent men as a class. "Understanding of history and the solution of social problems will be advanced as much by the study of eminent men as by the study of criminals, the unemployed and the mentally deficient."

Author Index to Volume 7

In this index are listed the names of the authors and the titles of their articles. Abstracts of papers read at meetings and articles that appear in the "Digest of Periodical Literature" section of the journal are not listed; they are indexed in the Analytic Subject Index.

Erratum. In the note by J. K. Robertson, page 259, line 8, substitute "polarizer" for "analyzer."

Analytic Subject Index to Volume 7

The titles of articles are disregarded, the entries being based on analyses of the contents of the original articles. Entries marked (D) refer to digests appearing under "Digests of Periodical Literature" and to abstracts of papers read at meetings; entries marked (R) refer to reviews appearing under "Recent Publications and Teaching Aids."

To facilitate reference to any desired subject, the index is divided into sections arranged alphabetically. The titles of these sections and of various cross-references to them are as follows:

nd
H.
330

nd

196
nd

t—

nd

re-

ws)
en

M.
O.
of
ion
in
. J
der
of
T.
nd

win

fied

rd,

185

D.
nd
on,

24,